



## Evaluation Design Options Report

## Career Pathways Design Study

Contract # DOLQ129633231  
/DOL-OPS-16-U-00055

February 2018

*Prepared for:*  
U.S. Department of Labor  
Chief Evaluation Office  
200 Constitution Avenue, NW  
Washington, DC 20210

*Submitted by:*  
Laura R. Peck  
Matthew Zeidenberg  
Sung-Woo Cho  
Daniel Litwok  
Julie Strawn  
Maureen Sarna  
Karin Martinson  
Deena Schwartz

**Abt Associates**  
4550 Montgomery Avenue  
Suite 800 North  
Bethesda, MD 20814

This report was prepared for the U.S. Department of Labor (DOL), Chief Evaluation Office by Abt Associates, under contract number DOLQ129633231. The views expressed are those of the authors and should not be attributed to DOL, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.

## **Acknowledgements**

The authors of this report are grateful for input from Megan Lizik, Michelle Ennis, and Molly Irwin at the U.S. Department of Labor; and from the project's technical working group, which included the following members: Caitlin McLean (Center for the Study of Child Care Employment, University of California, Berkeley), Jennifer Craft Morgan (The Gerontology Institute, Georgia State University), Marc Scott (Steinhardt School of Culture, Education and Human Development, New York University), Jeff Smith (Professor of Economics, Department of Economics, University of Wisconsin-Madison), Lindsay Page (School of Education, University of Pittsburgh). Bry Pollack provided editorial assistance.

**Table of Contents**

**Acknowledgements** ..... **i**

**Table of Contents**..... **ii**

**List of Exhibits**..... **iii**

**1. About the Career Pathways Design Study** ..... **1**

    1.1 Findings from the Knowledge Development Phase ..... 2

    1.2 Research Questions and the Organization of This Report..... 7

**2. Impact on Participant Outcomes** ..... **10**

    2.1 Introduction to Impact Research Questions..... 10

    2.2 Research Approaches ..... 11

    2.3 Data Sources..... 25

    2.4 Practical Considerations and Tradeoffs..... 29

**3. Workers’ Career Trajectories in the Economy** ..... **32**

    3.1 Introduction to Career Trajectories Research Questions ..... 32

    3.2 Research Approaches ..... 33

    3.3 Data Sources..... 38

    3.4 Practical Considerations and Tradeoffs..... 42

**4. Role of the Public Workforce System**..... **43**

    4.1 Introduction to Workforce System Research Question ..... 43

    4.2 Research Approaches ..... 44

    4.3 Data Sources..... 45

    4.4 Summary ..... 46

**5. Cost Analyses** ..... **47**

    5.1 Introduction to Cost Research Questions ..... 47

    5.2 Research Approaches ..... 48

    5.3 Data Sources..... 53

    5.4 Practical Considerations and Tradeoffs..... 55

**6. Summary and Framework for Considering Research Options** ..... **56**

    6.1 What to Ask: The Causal or Descriptive Nature of a Research Question ..... 56

    6.2 How to Ask: Approaches to Research Questions and Framework for Considering  
Options ..... 57

    6.3 Where to Ask: Location Considerations..... 60

    6.4 Conclusion..... 61

**References** ..... **62**

List of Exhibits

Exhibit 1. Four Sets of Career Pathways Research Questions (as a Map to This Report) ..... 8

Exhibit 2. Example of 3x3 Factorial Design ..... 16

Exhibit 3. Illustrative Minimum Detectable Effect Size Implications of Alternative Evaluation Designs..... 20

Exhibit 4. Example – Sankey Diagram of Occupational Transitions of Nursing, Psychiatric, and Home Health Aides..... 35

Exhibit 5. Expected Costs of a Career Pathways Intervention, by Perspective..... 50

Exhibit 6. Cost Analyses in Completed and Ongoing Studies ..... 55

Exhibit 7. Framework for Considering Tradeoffs across Research Approach Options for Causal Questions ..... 58

## 1. About the Career Pathways Design Study

Career pathways approaches to workforce development offer articulated education and training steps between occupations in an industry sector, combined with support services, to enable individuals to enter and exit training within a pathway at various levels and to advance over time to higher skills, recognized credentials, and better jobs with higher pay. Each step on a career pathway is designed explicitly to prepare individuals to progress to the next level of employment and/or education. Career pathways strategies have a sectoral focus, targeting jobs in industries of importance to local and regional economies and striving to build strong relationships with employers. States and localities across the United States have increasingly adopted career pathways approaches.<sup>1</sup>

The career pathways framework evolved over the last decade as a response to emerging evidence on labor market changes and on the limits of previous employment and training strategies. In the labor market, individuals with a high school education or less experienced stagnating wages and relatively high unemployment over the last 30 years, whereas those with postsecondary credentials experienced economic gains (Autor, 2015; Carnevale, Jayasundera, & Gulish, 2016). In the workforce development field, researchers studying long-term outcomes found that the two most common employment strategies for low-income adults—quick job placement or stand-alone basic skills instruction—neither increased employment and earnings over the long run nor helped participants escape poverty (Hendra & Hamilton, 2015). By emphasizing postsecondary job skills, career pathways approaches seek to respond to these labor market changes and to deliver larger and longer-lasting impacts than previous employment and training strategies. The career pathways framework also seeks to incorporate promising features of recent workforce development innovations, such as targeting industry sectors and integrating basic education with job training (Werner et al., 2013). In addition, the career pathways approach involves providing a range of supports to students including advising, financial assistance, and connections to the labor market and jobs.

The rapid rise of career pathways strategies, including an emphasis on them in the Workforce Innovation and Opportunity Act (WIOA), creates a need for more evidence on this approach. To date, research on career pathways programs suggests that they may prove more effective than earlier generations of employment and training strategies if early positive impacts on attainment of educational credentials (both certificates and degrees) ultimately translate into sustained economic benefits (Elliott & Roder, 2017; Rolston, Copson, & Gardiner, 2017; Martinson, Copson, Gardiner, & Kitrosser, 2018). Although substantial career pathways research is in progress, more research is needed to enhance the field’s knowledge about career pathways strategies, how best to configure them, and their long-term effects.

WIOA requires the U.S. Department of Labor (DOL) to “conduct a multistate study to develop, implement, and build upon career advancement models and practices for low-wage healthcare providers or providers of early education and child care” (29 U.S. Code § 3224(b)(4)(I)). In response, the Chief Evaluation Office (CEO) at DOL funded the Career Pathways Design Study and contracted

<sup>1</sup> For a comprehensive history of how the career pathways framework evolved, see *The Evolution and Potential of Career Pathways* (U.S. Department of Education, 2015). For a review of how career pathways program and system approaches are currently defined in national frameworks and in practice, see this project’s *Career Pathways Implementation Synthesis* report (Sarna & Strawn, 2018).

with Abt Associates to develop evaluation design options that could address critical gaps in knowledge related to the approach, implementation, and success of career pathways strategies generally, and in early care and education specifically (given the scarcity of information on it relative to healthcare).

This report provides information on potential research approaches to answering key questions about career pathways strategies that we identified as most relevant to the workforce development field and that also are less likely to be fully addressed by current ongoing research. It outlines a range of design options informed by completed and ongoing studies. These options also reflect the current state of implementation of career pathways program- and system-level initiatives, as well as respond to DOL's interest in improving employment and earnings outcomes for youth and adults.

Before detailing these research approaches, this report summarizes what we learned in the first, knowledge development phase of this project about current evidence and implementation in the career pathways field. It then describes priority research questions that emerged from that work and concludes by outlining the remainder of the report.

### 1.1 Findings from the Knowledge Development Phase

To understand the current state of the career pathways field and inform the evaluation design options to be presented in this report, Abt first undertook a knowledge development effort that produced reports on (1) research and evaluation relevant to career pathways approaches, (2) the implementation of existing career pathways initiatives, and (3) the potential for career pathways approaches in early care and education (ECE). This section summarizes the scope, methodologies, and major findings from these reports. Because the approaches used in the first two reports overlap, we summarize them together below, and then describe the ECE report separately.

#### 1.1.1 Career Pathways Research/Evaluation and Implementation

##### Scope and Methodologies

The first report (*Career Pathways Research and Evaluation Synthesis*; Schwartz, Strawn, & Sarna, 2018) explored relevant research and evaluation. This review focused primarily on the type, scope, and setting of career pathways research that either has been completed or is ongoing; on what research questions have been asked; key findings to date; and areas for further research.

We conducted a broad scan of research on career pathways interventions as of February 2017, which identified 52 studies, both complete and ongoing. We included studies of interventions that (1) focus on adults (including young adults, but excluding high school students); (2) include occupational training; and (3) describe themselves as involving career pathways, or which include at least some key element(s) of the pathways approach.<sup>2</sup>

The research report categorized studies as examining either *system-level* or *program-level* interventions, or both. System-level interventions address, at least in part, the six career pathways systems elements that DOL described in its Career Pathways Toolkit (2015). The six elements are (1) build cross-agency partnerships and clarify roles, (2) identify industry sectors and engage employers,

---

<sup>2</sup> For example, we included studies of several training efforts that were not explicitly career pathways approaches—in some cases predating the concept—where there were major studies of interventions representing core elements of the career pathways model, such as sector training partnerships or integrated basic education and training.

(3) design education and training programs, (4) identify funding needs and sources, (5) align policies and programs, and (6) measure system change and performance.

We categorized interventions as program-level if they included elements of the WIOA definition of career pathways: programs that provide individualized training and supports that (1) align with the skill demands of the state and local economy; (2) prepare individuals to be successful in a range of secondary and postsecondary education options; (3) include academic and career counseling, as well as non-academic supports; (4) provide, as appropriate, concurrent and accelerated program designs; and (5) help individuals to enter or advance within a specific occupation or occupational cluster.

Many of the interventions documented in the report included only some of the elements in these system-level and program-level definitions.

Whether system or program level, we also noted in our analysis whether or not an intervention provided more than one step of education or training in a career pathway; or if not, that it at least closely linked its training to the specific next training step(s) and actively helped participants enroll in that next step. We chose to highlight multiple steps of training as a proxy for understanding the extent to which sites in a study focused on career advancement, arguably the most distinctive feature of the career pathways framework. By “steps” we mean training opportunities that result in successively higher levels of skills within a career pathway, such as entry-level, mid-level, and advanced training, which correspond to specific, increasingly better-paying jobs.

The second report (*Career Pathways Implementation Synthesis*; Sarna & Strawn, 2018) synthesized high-level information about implementation of 128 career pathways initiatives as of February 2017. Drawing on information from publications and websites, as well as insights from 21 experts with experience at the local, state, and national levels in career pathways approaches, it described the type, scope, and setting of past and present career pathways initiatives; described how career pathways approaches appear to be defined in practice; summarized experts’ views about the state of the career pathways field, successes and challenges, and priority research needs; and outlined possible directions for future research.

### **Findings in Brief<sup>3</sup>**

The *Research and Evaluation Synthesis* report identified 52 studies of career pathways approaches and related efforts (such as sectoral training). We used the presence of multiple steps of training as a proxy for how “fully developed” a career pathways effort was. Through this lens, about three-quarters of studies (39 of 52) focused on career advancement involving multiple steps of training (and had at least one site that offered multiple steps of training). That is, they were “fully developed.” In just 12 of these 39 studies, all of the sites studied met that definition. This suggests that the research on fully developed career pathways initiatives is fairly limited.

In contrast, in the *Implementation Synthesis* report, which looked at career pathways efforts regardless of whether they were the subject of research or not, a higher proportion of initiatives (54 of 128, or 42 percent) consisted of efforts in which every site included multiple steps of training. This suggests that

---

<sup>3</sup> For a complete discussion of findings, including matrices with details of each study and implemented initiative reviewed, please consult the full reports: Schwartz et al. (2018) and Sarna & Strawn (2018).

there may be a greater focus on pathway advancement out in the field than is captured in career pathways research at this time.

We also found that program-level interventions are much more numerous than system-level ones. Among the 52 research studies, 51 of them looked at program-level efforts; and, although 20 looked at system-level interventions, the research on most of them was in conjunction with program-level interventions. However, these numbers may understate the reach of system-level efforts because they typically have a broader geographic scope (e.g., an entire state or region). Just two studies explicitly conducted systems-change analyses. Among the career pathways efforts in practice, about one-fifth (23 of 128) of initiatives were solely system-level efforts. This suggests that there may be greater focus on implementing system-level career pathways efforts in the field than is reflected in the research.

Career pathways efforts in the healthcare sector were by far the most commonly implemented (and researched). Other common sectors included manufacturing, information technology, and business (e.g., accounting, office/clerical occupations). Career pathways programs most commonly targeted low-income individuals; typical participants were high school graduates in their late 20s or 30s and more likely to be female than male.<sup>4</sup> Few studies included substantial percentages of people of Hispanic descent or youth, and few included significant percentages of individuals with employment barriers such as no high school diploma, very low skills (less than eighth grade), limited English proficiency, criminal records, or disabilities.

The studies asked a diverse set of research questions, most commonly focused on what the intervention was, how it was implemented, and whether it was effective. Thirty-eight of the 42 studies for which we identified outcomes included at least one employment outcome (most frequently earnings, employment, and hourly wage); 29 included at least one education outcome (most frequently the receipt of a certificate or degree). Only about one-third of the studies in our review included research questions that either specifically mentioned career pathways in a question or included questions about an intervention that was explicitly described as a career pathways approach.

Among the studies we identified, findings have been published for four *quasi-experimental* and eight *experimental* evaluations—using methods that allow researchers to attribute outcomes to the program—that measured the impact of the program on participants’ outcomes.<sup>5</sup> (This is in contrast to *non-experimental*, descriptive studies, which document program implementation or measure and describe participants’ outcomes, but do not use a comparison or control group to approximate what would have happened in the absence of the program and therefore determine whether the program caused a change in outcomes.) Highlights of the findings include the following:

---

<sup>4</sup> Career pathways participants may be more likely to be female because many studies were of the healthcare sector, which tends to employ more women than men.

<sup>5</sup> In brief, an experimental evaluation involves randomly assigning eligible units (be they people, organizations, or geographic locations, for example) to gain access to one condition (a policy, program, or intervention) or to be denied such access (control). Any differences between outcomes for the two groups are attributable to the intervention under study. A quasi-experimental design uses methods other than random assignment to estimate a counterfactual condition (or what would have happened in the absence of the intervention), from which to estimate impacts.



- Ten of these published impact findings are for short-term (one to two years) and medium-term (three to four years) follow-up periods. Two studies to date have reported impact findings for long-term outcomes (five years or longer). This represents a significant evidence gap because analyses comparing short-term and long-term impacts for education and training programs find that results can change substantially over time, with economic impacts for skill development often small initially but growing larger with longer-term follow-up (Card, Kluve, & Weber, 2010; King, 2004). In addition, short- and medium-term follow-up likely will fail to capture the extent to which participants advance or not through more than one step of career pathways training and/or employment over time—a central question for the career pathways model.
- Nine of these impact studies examined earnings. Three found statistically significant positive results, five found mixed results, and one found mostly negative results.<sup>6</sup>
- Ten of the studies examined educational outcomes. Seven found statistically significant positive results, one found mixed results, and two found mostly negative results.
- Of the eight random assignment studies reporting impacts, only one included at least one site that offered multiple steps of training.
- Across all of the studies we examined, researchers found that implementing a model as intended often proved challenging. In addition, programs varied considerably in the populations targeted and served, and in targeted sectors and occupations.

Looking forward, more than half of the research studies on career pathways initiatives are ongoing studies, and more than half of all the implemented career pathways initiatives we examined are still active. Ten of the studies will report impact results by 2018 and 26 of them will report results by 2021; 76 initiatives currently operate in the field. This means that in the next five years, considerably more will be known about career pathways strategies both from research and from practice.

As described below, these future career pathways reports will fill some specific gaps in the existing evidence, and will do so using more rigorous research than many of the previous studies. More than twice as many ongoing studies use random assignment and quasi-experimental methods (24) as did completed studies (11). Important questions that will be addressed by future research include these:

- *Research explicitly focused on the career pathways model.* Of the studies that included research questions explicitly focused on career pathways, 61 percent are ongoing. This includes five of the six experimental studies with explicit career pathways questions, and two studies aimed at understanding how career pathway programs and efforts can and do use metrics to understand those pathways in practice.<sup>7</sup>

---

<sup>6</sup> What we call “mixed results” include two or more of the following: statistically significant positive results, statistically significant negative results, or non-statistically significant positive or negative results.

<sup>7</sup> The Alliance for Quality Career Pathways is working with state career pathways efforts to determine the predictive power of interim metrics. The Career Pathways Programming for Lower-Skilled Adults and Immigrant study is a descriptive study addressing what short- and long-term outcome measures adult education providers commonly use and report for career pathways participants.

- *Research on career pathways approaches offering more than one step of training.* More ongoing studies (22) include at least one site with multiple steps of career pathways training than did completed studies (17). While the most common occupational training provided in these studies is for entry-level positions, our *Research and Evaluation Synthesis* report suggests many sites included in ongoing studies also offer preparation for mid-level jobs.
- *Long-term outcome and impact findings.* A number of the ongoing studies are following participants for five years or more and will be able to report on long-term outcomes and impacts on employment and earnings.
- *More analysis of program costs.* Ten of the ongoing studies are examining program costs but only four were completed studies. Some of these also are examining questions of program sustainability.
- *Exploration of the effectiveness of specific components within a career pathways approach.* Three of the ongoing studies have research questions about the impact of specific program components, such as the Health Profession Opportunity Grants (HPOG) Impact Study's test of facilitated peer support, emergency assistance, and non-cash incentives within a career pathways model.
- *More analysis of systems change initiatives.* Among the 20 studies that analyze career pathways system change initiatives, 13 are ongoing as compared with just seven completed.<sup>8</sup>

### 1.1.2 Career Pathways Initiatives in Early Care and Education

#### Scope and Methodology

The third report (*Career Pathways in Early Care and Education*; Cheng et al., 2018) examines the potential for career pathways approaches in the ECE sector, which includes child care, preschool, and prekindergarten services. Consistent with the WIOA legislative mandate for this project, it focuses on low-wage workers in jobs providing direct care and teaching in the ECE field. For the analysis in this report, we reviewed current research about the ECE workforce released by federal agencies and prominent organizations engaged in ECE research, and we held group discussions with 23 experts from federal and state agencies, Head Start and prekindergarten systems and programs, community colleges, leaders of national ECE organizations, and researchers on ECE.

#### Findings in Brief

The *Early Care and Education* report identified few initiatives aimed at creating ECE career pathways strategies as of February 2017. Among research studies on career pathways initiatives and among implemented career pathways efforts, about 10 percent included ECE occupations. The main career pathways program models that exist in this sector are Registered Apprenticeship and the Teacher Education and Compensation Helps (T.E.A.C.H.) Early Childhood<sup>®</sup> Initiative.

Rather than comprehensive career pathway approaches, the report found promising practices aimed at addressing various barriers specific to ECE worker advancement, such as wage and compensation supplements, and efforts to provide portable credentials and articulation and credit for prior learning.

---

<sup>8</sup> These studies are addressing questions about how various institutional types implement career pathways approaches, the sustainability of career pathways approaches, and how different types of partners contribute to career pathways approaches.

These practices hold the possibility for becoming the building blocks of comprehensive career pathways in the ECE field, but given pervasive low wages in the field (outside of public school employment) it is unclear whether the structure of employment and compensation in the ECE field offers enough advancement opportunities to make career pathways strategies viable.

### 1.2 Research Questions and the Organization of This Report

The three reports detailed above each culminated in the description of a number of research questions that, if answered, would advance our understanding of career pathways system- and program-level strategies. We narrowed the questions in collaboration with DOL to a list of research questions that reflect those most relevant to the workforce development field and that also are less likely to be fully addressed by current ongoing research.<sup>9</sup>

These questions, shown in Exhibit 1, are organized into four groups, which provide the organizational structure for this *Evaluation Design Options* report: program impact on participant outcomes; workers' career trajectories in the economy;<sup>10</sup> the role of the public workforce system in career pathways initiatives; and cost analyses. That is, each set of questions is discussed in detail in the four report chapters that follow.

- Chapter 2—We examine three research questions that explore the *impacts on participant outcomes* of career pathways approaches. We consider the *long-term impacts* of career pathways programs as a whole, impacts of *components* of these programs, and impacts on *subgroups* of interest.
- Chapter 3—We examine questions about *career trajectories* that consider descriptively how workers generally (not just workers in career pathways programs) progress through education, training, and employment, and specifically how those experiences occur in the ECE sector.
- Chapter 4—One research question about the *public workforce system* considers the role of that system in creating a supportive environment for career pathways efforts to be successful.
- Chapter 5—Finally, research questions about the *costs* of career pathways programs consider the balance of costs and benefits and the likely returns on investing in career pathways initiatives.

---

<sup>9</sup> As noted earlier, career pathways research currently underway will ultimately include more long-term outcome and impact findings than are available now, as well as findings on impacts of certain career pathways components and subgroups. However, we expect that there will still be additional questions and a need to bolster the evidence base in these areas. As discussed in Chapter 2, some possible approaches would build on the work of these ongoing studies.

<sup>10</sup> The purpose of looking descriptively at workers' career trajectories in the economy would be to set the context and expectations for what might be achieved in career pathways programs based on the experience of workers in the labor market more generally (not just those in career pathways programs).

**Exhibit 1. Four Sets of Career Pathways Research Questions (as a Map to This Report)**

Set	Research Questions	Possible Research Approach(es)
Impact on participant outcomes (Chapter 2)	<ul style="list-style-type: none"> <li>• <b>Long-term Impacts:</b> What impacts do career pathways programs have on participants' advancement through multiple, progressively higher levels of education and training, and associated jobs, within a pathway over time? To what extent do participants move beyond entry-level training and employment, even over a long follow-up period?</li> <li>• <b>Component Impacts:</b> Which components of career pathways programs are the strongest drivers of impacts? To what extent is the impact of a combination of career pathways components greater than that of any individual part? That is, are certain components more effective when bundled together?</li> <li>• <b>Subgroup Impacts:</b> For what groups are career pathways programs most effective? To what extent can career pathways approaches influence groups not commonly included in career pathways research to date? And which components matter most for which subgroups?</li> </ul>	<ul style="list-style-type: none"> <li>• Descriptive analyses</li> <li>• Meta-analysis</li> <li>• Build on existing impact evaluations</li> <li>• Design new (experimental) impact evaluations</li> <li>• Use quasi-experimental designs                             <ul style="list-style-type: none"> <li>- Comparison group</li> <li>- Time series</li> <li>- Regression discontinuity</li> </ul> </li> <li>• Use quasi-experimental analytic approaches                             <ul style="list-style-type: none"> <li>- Difference-in-differences</li> <li>- Individual fixed effects</li> <li>- Cross-state analyses</li> </ul> </li> </ul>
Workers' career trajectories in the economy (Chapter 3)	<ul style="list-style-type: none"> <li>• In the absence of career pathways programs, to what extent, and how, do workers advance on their own through multiple, progressively higher levels of education and training, and associated jobs, within a pathway over time? To what extent is advancement more common for workers from specific settings or backgrounds? To what extent do wages increase for workers who progress through these pathways? What are the implications for designing and evaluating career pathways initiatives?</li> </ul>	<ul style="list-style-type: none"> <li>• Use extant survey and administrative data to identify trajectories that occur commonly in the labor market</li> <li>• Design a new survey instrument to track workers, using a sample that is representative of the labor market as a whole or a single industry sector</li> <li>• Use predictive analytics to determine possible career trajectories of workers</li> <li>• Use qualitative research methods to better understand career trajectories</li> </ul>
The role of the public workforce system in career pathways initiatives (Chapter 4)	<ul style="list-style-type: none"> <li>• What roles in career pathways initiatives is the workforce system playing as compared with other entities that more commonly lead these strategies, such as community and technical colleges?</li> </ul>	<ul style="list-style-type: none"> <li>• Review existing studies</li> <li>• Conduct implementation research</li> <li>• Conduct pilot studies</li> </ul>
Cost analyses (Chapter 5)	<ul style="list-style-type: none"> <li>• What are the costs associated with a given career pathways program to participants, funders, and employers, including opportunity costs? What do those costs "buy"? And to what extent are the benefits worth the cost?</li> </ul>	<ul style="list-style-type: none"> <li>• Conduct cost allocation analysis, cost-effectiveness analysis, or cost-benefit analysis. Each of these represents different strategies for monetizing programs/initiatives' costs and/or benefits such that they can be compared to each other or to other programs/initiatives.</li> </ul>

For each set of research questions, the report describes potential research approaches, possible data sources for research, and practical considerations and tradeoffs. We describe what we mean by each of these terms below.

- A **research approach** refers to how one might go about answering the research question. An approach can follow any line of thought or use any method. For each set of research questions, we elaborate on the varied, relevant possible approaches. We describe the approaches in general terms; a detailed analysis plan should be developed before implementing any of these.

- ***Data sources*** refers both to existing impact evaluation data for the completed or ongoing studies identified in our knowledge development work and to other data not tied to a particular evaluation, such as the Current Population Survey. For each research question, a single data source or a combination of data sources might be relevant.
- ***Practical considerations and tradeoffs*** refers to issues related to implementation, timing, and cost of each of the proposed approaches and data sources. For each set of questions, we highlight relevant issues that should be considered in deciding whether and how to move forward with the associated research.

After exploring each set of research questions across these three dimensions in-depth in Chapters 2–5, in Chapter 6 we address issues of what locations are fitting for research and also propose a framework for considering the practical considerations and tradeoffs across the sets. We conclude with options for how DOL could prioritize and plan for career pathways research in the future.

## 2. Impact on Participant Outcomes

This chapter focuses on the subset of the priority research questions that relate to estimating the impact of career pathways programs on participants' outcomes, either in whole programs or their component parts, including impacts on subgroups of interest. After introducing and discussing the questions in Section 2.1, Section 2.2 discusses five kinds of approaches for answering these questions. Section 2.3 discusses relevant data sets, and Section 2.4 discusses practical considerations and tradeoffs.

### 2.1 Introduction to Impact Research Questions

Three of the Career Pathways Design Study's priority research questions focus on the impacts of career pathways programs on participants' outcomes. We label them as being about long-term impacts, component impacts, and subgroup impacts, as follows:

**Long-term Impacts:** What impacts do career pathways programs have on participants' advancement through multiple, progressively higher levels of education and training, and associated jobs, within a pathway over time? To what extent do participants move beyond entry-level training and employment, even over a long follow-up period?

**Component Impacts:** Which components of career pathways programs are the strongest drivers of impacts, especially on credential attainment, employment, and earnings? To what extent is the impact of a combination of career pathways components greater than that of any individual part? That is, are certain components more effective when bundled together?

**Subgroup Impacts:** For what groups are career pathways programs most effective? To what extent can career pathways approaches influence groups not commonly included in career pathways research to date? And which components matter most for which subgroups?

These priority research questions are aimed at filling gaps in the current career pathways evidence base identified in the knowledge development phase of this project. As noted earlier, substantial completed research and research underway (Schwartz et al., 2018) is poised to answer questions of overall or average impacts in the short term (one to two years) and medium term (three to four years). Some studies underway will also include long-term follow-up, but given the lack of long-term impact findings (five years or longer of follow-up) to date, and the importance of long-term impacts to the field, we expect these questions will remain a priority. Beyond these long-term impacts, the priority research questions also address similar knowledge gaps concerning the *what works* (components) and *for whom* (subgroups) aspects of career pathways program impacts. While ongoing career pathways research will ultimately yield findings on certain components and subgroups, we expect there still to be a need for additional evidence in these areas, too. There are several considerations in each of the three areas of study for the impact research questions.

The *Long-term Impacts* questions ask about the extent to which there is evidence for the underlying theory of career pathways approaches in individual participants' actual and observed experiences. Based on this theory, there are progressions of education, training, and work experiences that build on one another to advance people along a career path. However, the existing research to date has not observed these over a period long enough to establish support for the theory. Empirical evidence also

points to examining long-term impacts: because the full employment and earnings impacts of skill development can take time to materialize, short- and medium-term (one to four years) impacts look different from longer-term impacts (five to 10 years). Current research typically observes first jobs after training, without (yet) enough follow-up to understand when and how people return to education, access additional training, and use those to propel themselves into better jobs along a career path.

To address the *Long-term Impact* research question requires two things: time and attention to measurement over time. First, future research should consider how to measure what is not a simple construct: workers' education and employment histories. Whereas it is relatively straightforward to track employment, earnings, education, and training at a single point in time, it is more complex to compile a history of those as they intertwine and to understand how these experiences overlap or which affects the next. For example, how does the length or type of initial training affect whether an individual subsequently progresses to a higher level of training and employment? How do these experiences influence additional training or education and subsequent jobs, and their quality?

In considering the *Component Impacts* questions more fully, we identified some specific program components that appear to be of interest to the field as focus areas.<sup>11</sup> These include instructional and curricular innovation (including use of technology), integrated basic education and training, career coaching, academic supports, financial aid, support services (child care, transportation, etc.), job development, work-based learning, employment retention services, and training stipends.

Finally, the *Subgroup Impacts* questions focus on what works best for whom. This includes research to inform how career pathways programs can be designed and operated to serve groups underrepresented in studies to date (e.g., individuals of Hispanic descent, youth) and groups with specific employment barriers such as no high school diploma, very low skills (e.g., less than eighth-grade literacy or numeracy), limited English proficiency, criminal records, or disabilities.

## 2.2 Research Approaches

This subsection discusses five kinds of research approaches for answering questions about long-term, component, and subgroup impacts: (1) conducting descriptive analyses; (2) conducting meta-analyses that aggregate impact estimates across studies to provide an overall impact estimate; (3) building on existing evaluations; (4) undertaking new experimental evaluations; and (5) conducting quasi-experimental analyses.

### 2.2.1 Descriptive Analyses

The most rigorous evidence on the effects of career pathways programs on participant outcomes will come from experimental evaluation research. Descriptive research—which relies on such methods and sources as surveys, interviews, observations, focus groups, and administrative data analyses—can play a critical role in career pathways research, too, by documenting *what* exactly is being implemented, *how* it is being implemented, how implementation *changes* over time, and in what *context*. It can also provide insights about stakeholder perspectives, participant experiences, program operations, and program outcomes (though not about what caused those outcomes). Descriptive research in this setting has four main goals: (1) providing a comprehensive description of each career

---

<sup>11</sup> See *Career Pathways Implementation Synthesis* for a summary of our conversations with experts in the field.

pathways program and/or system initiative being studied; (2) informing impact questions, particularly as related to understanding the effects of specific career pathways components and the bundling of those components; (3) creating coded measures of program design and implementation strategies, which can be used for analyses of component impacts as part of pooling/analyzing existing impact evaluations; and (4) identifying areas for future inquiry.

The first of these goals (providing a comprehensive description of implementation) is important in its own right, such as for determining fidelity to the career pathways model being tested and for considering how a given program is like or unlike other programs (as needed for making claims about the generalizability of results, for example). Most generally, research describing what a program is draws on implementation evaluation techniques. Other forms of descriptive research also hold promise for being useful in conducting research to inform impact questions. For example, collecting and analyzing data about the perspectives of key stakeholders and experiences of participants in a given program may help not only in interpreting impact evaluation results but also in identifying and prioritizing specific impact questions to ask and where to study them.

Descriptive evidence that comes as part of an experimental or quasi-experimental evaluation provides detail on what services are provided both to participants who could access the intervention (“treatment group”) and to participants who could not (“control group”); this produces information on the extent of service differential. It is these differences in access that are the likely drivers of any subsequent impacts observed. Moreover, understanding the duration and content of services, particularly the key elements of the career pathways approach, is important both for research interpretation and for policymakers and program administrators. Analyses of experimental data can assess the impact both of access to and receipt of services.

Later we discuss how pooling across multiple studies—pooling individual-level data sets or pooling aggregate-level study data—holds promise for understanding the impacts of selected program components and implementation features. In order to do this, consistent program-level measures must exist. Qualitative research can be useful in creating those coded measures, ensuring that they represent the aspects of program design and implementation that are important for analysis. This is relevant for cross-site analyses of existing evaluations, as well as for meta-analyses (discussed next).

Descriptive research also can be useful in helping to identify which program components can and *should* be part of more rigorous future evaluation. For example, surveys, interviews, and/or focus groups of program participants, program staff, and/or administrators can help provide information on these stakeholders’ perceptions of component effectiveness. These varied perspectives can be helpful in determining what about *program configurations* is perceived to be the most helpful. Program participants might report that one-on-one counseling was essential to their success; program staff might report that financial assistance is critical; and program administrators might point to college remediation and/or Adult Basic Education as important to their programs’ success. These insights would help researchers identify—at least qualitatively—which program components are perceived to be the most effective and therefore might be targets for future, rigorous testing, either alone or in specific combination with other components.

### 2.2.2 Meta-Analyses

A meta-analysis involves systematically analyzing the results from existing studies. It does so by aggregating impact estimates across studies to provide an overall impact estimate. Meta-analyses also can ascertain something about the relative effectiveness of various program designs and/or also about



how impacts might vary across subgroups. A meta-analysis first composes a sampling frame, identifying the criteria under which studies would be included in the analysis. In this case, the studies would pertain to the effectiveness of career pathways strategies. From there, a meta-analysis imposes some criteria on the study's design, usually prioritizing high-quality designs for inclusion in the analysis, and outcome measures, ensuring that there is some comparability across outcomes examined. A meta-analysis then will code various aspects of the research and the intervention; and this coding permits focus on particular program components that are part of the intervention and particular subgroups, reporting results for both. Indeed, a strength of meta-analysis for the purpose of this research would be to examine variation across studies to shed light on certain strategies or populations.

The *Research and Evaluation Synthesis* report (Schwartz et al., 2018) provides a solid sample (of 52 studies) that might be considered as a starting point for meta-analyses. To refine that list further, this report (see Appendix A, Exhibit A.1) provides a list of studies (both experimental and quasi-experimental) that could be part of a meta-analysis of the *impact* of particular components of career pathways programs or of the *impact* of career pathways for particular subgroups. These are the experimental and quasi-experimental impact studies that could be judged to determine whether they meet a “quality” criterion for meta-analysis.<sup>12</sup> If so, a meta-analysis of these programs could focus on specific program components and subgroups, in the interest of informing the priority research questions. It is possible that additional studies could be included as part of a meta-analysis, but they might be given lower weights as less relevant in instances where their evaluation designs are weaker.

### 2.2.3 Build on Existing Impact Evaluations

Perhaps the strategy that holds the greatest promise—at the lowest investment—for addressing the priority research questions is that of tapping existing impact evaluation data to learn more.<sup>13</sup> This can occur in two ways: re-analyzing existing data or conducting additional follow-up on existing evaluations' samples. We discuss both options, as appropriate to the impact priority research questions. Building on existing impact evaluations differs from the meta-analysis approach discussed above because meta-analysis considers study-level data, whereas this approach reanalyzes or adds to individual-level data.

#### Long-term Impacts

To assess the long-term impacts of career pathways strategies, research could involve longer-term follow-up on existing studies. Such follow-up could include conducting new primary data collection, through follow-up surveys of participants, for example, or linking to other (mostly administrative) data sources to analyze longer-term impacts on participants' education and employment trajectories. The main national source of data for long-term *educational* follow-up would be the National Student Clearinghouse, with state and local student-level data also important sources; the main sources of *employment* data for follow-up would be earnings records from the Internal Revenue Service (IRS), the Social Security Administration, state Unemployment Insurance (UI) programs, the Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program, and the U.S. Department of

---

<sup>12</sup> Rather than filtering by some quality criterion, all studies could be included in a meta-analysis that would then use covariates to account for study quality.

<sup>13</sup> As discussed in Section 2.4 this approach would require researchers to examine past informed consent language to determine whether further work with the sample is permissible, or if researchers would need to seek new consents.

Health and Human Services, Administration for Children and Families' National Directory of New Hires (NDNH). Section 2.3 on data sources provides additional detail on each of these (see Appendix A, Exhibit A.2).

### **Component Impacts**

Existing evaluation data can be useful as a source for information on the relative effectiveness of program components. A meta-analysis (discussed above) is one option; another option involves pooling the site-level and individual-level data across studies to consider which program components appear to influence program impacts. To permit this type of analysis, comparable measures of program components and implementation features must be available. These are available across some program evaluations; evaluations lacking them could be the subject of new data collection, permitting more evaluation data sets to be part of this kind of pooling and associated analysis.<sup>14</sup>

### **Subgroup Impacts**

The priority research question about how career pathways strategies function for various subgroups could be addressed by re-analyzing existing evaluation data to consider new subgroups (that were not part of the original studies' analysis), sample sizes permitting. One way to overcome sample size limitations in a given study would be to pool across multiple studies, creating sufficient sample within a subgroup of priority interest. In response, DOL should prioritize those subgroups that are of greatest interest, from which studies, and considering which outcomes.

Research on the *Component Impacts* and *Subgroup Impacts* questions can be brought together to consider which program components seem to be most or least important for selected subgroups.

### **2.2.4 Design New Impact Evaluations**

Existing studies—though providing a useful resource for career pathways research—are limited in the programs they focus on, the data they collect, and the outcomes they have selected for study. Should DOL use prior studies to help answer research questions of interest, some limitations will exist. For example, because the baseline data will already have been collected, subgroup identifiers are essentially fixed. Moreover, sample sizes—and subgroup sample sizes—were determined to support those prior studies' research questions and not the research questions of the future. As a result, the strategy that has strong potential to allow DOL to answer specific priority research questions without compromise is that of mounting new impact evaluations. Doing so would allow DOL the option to specify the research design to respond directly to the research questions of high priority.

For example, new impact evaluations could collect more detailed baseline information on previous employment and training experiences to help construct career profiles. In doing so, a longer-term picture of employment and training connections could be developed, not necessarily only by looking forward but also by looking retrospectively. In addition, new studies could oversample individuals with previous work experience and/or training in a targeted sector in order to analyze career advancement impacts for mid-career entrants versus new entrants to the sector. New studies might also test a narrower definition of career pathways strategies to focus on specific aspects of the approach.

---

<sup>14</sup> Prior analyses that have followed this strategy include Bloom, Hill, & Riccio's (2003) pooling of several welfare reform studies for analysis of what about the programs and their implementation associates with individual participant impacts. The HPOG Impact Study currently also is using this approach and will release findings in 2018.

All of these are aspects of new research that could focus on combinations of long-term impacts, program component impacts, and subgroup impacts of greatest interest.

### Long-term Impacts

In order to measure the long-term impacts of career pathways strategies, future studies can be designed with an informed consent that allows for longer-term follow-up on education and employment trajectories and related longer-term outcomes of interest. Even if additional long-term surveys are not part of a new study, the study could ensure that permission to match study participants' identifiers to administrative data at any future date is not restricted, so that researchers can perform all necessary long-term follow-up.

### Component Impacts

Future research could involve designing new studies specifically to isolate the effects of particular career pathways program components. These designs can involve more features than those of past evaluations, where the most common research design involved a single treatment and single control group. Future experimental evaluations could consider using multiple experimental arms with greater frequency than has been done in the past. Additional options include factorial designs, staged designs, adaptive and/or crossover designs, and cluster-randomized designs. Each of these types of designs is summarized as follows (and, thereafter, we provide a simple power analysis that shows how each of these more complex designs requires additional sample—relative to a two-group experiment—to detect the same impact magnitude):

- **Multi-armed experiments.** A multi-armed experimental evaluation design randomizes study participants to three or more groups (“arms”), where those groups can include a control group and multiple variants of a treatment, to ascertain which is preferable. In the case of seeking to understand which program components are essential ingredients of a program’s overall recipe, the “enhanced treatment design” (see Bell & Peck, 2016; Peck, 2015) is most relevant. An enhanced treatment design involves selecting a specific program component to either scale up or scale back in order to judge whether that specific component meaningfully improves or contributes to the program’s overall impact.

Example. In the career pathways context, participants could be randomized to any one of several program components for evaluation. For example, career coaching, financial aid, selected support services, work-based learning, job development, or training stipends could be provided to a randomly selected subset of program participants in order to judge its relative effectiveness on top of the standard program. That said, for the design to work, program administrators would need to be willing to withhold access to the component to any but the designated subset of participants. This means that some components might not be suitable for randomization—those judged as essential to the program (such as the core training, case management, transportation aid, or basic education requirements).

- **Factorial designs.** In a factorial design, selected aspects of a treatment (“factors”), each of which has two (or more) levels, are randomized separately. As Collins, Murphy, Nair, and Strecher (2005, pp.65-66) elaborate, a factorial design can aid in answering important questions: “which program components are working well; which should be discarded, revised, or replaced; which dosages of program components are most appropriate; whether delivery components are enhancing, maintaining, or diluting intervention efficacy; and whether individual and group

characteristics interact with program or delivery components.” Such questions usually are more important than the average treatment effect. A 2×2 factorial design could accurately be considered a four-armed experiment, where selected program components are randomized to judge their effectiveness both independently and jointly. In addition to absence/presence questions, the factorial design can accommodate randomizing to varying *levels* of a program component to identify whether more or less of it makes a difference.

Example. Consider two of the program components being evaluated (via a multi-armed experiment) as part of the Health Profession Opportunity Grants evaluation: emergency assistance and facilitated peer support. In HPOG, emergency assistance provides up to \$1,000 to help participants respond to emergencies, such as imminent eviction or work/school-necessary car repair. Facilitated peer support uses a cohort model to provide motivation, skills building, and accountability to participants. To test the *levels* (rather than simple presence/absence) of these two program components, a 3×3 factorial design could randomize participants to one of nine treatment groups, as illustrated in Exhibit 2.

**Exhibit 2. Example of 3x3 Factorial Design**

Emergency Assistance amount:	Peer Support frequency:		
	None	Weekly	Monthly
None	Standard Program	T2	T3
\$1,000	T4	T5	T6
Unlimited for 2 years	T7	T8	T9

This design would provide information on the extent to which emergency assistance interacts with participating in various levels of facilitated peer support. This design involves nine cells, where two cells (T2 and T3) consider the impacts of peer support frequency; two cells (T4 and T7) consider the impacts of amounts of emergency assistance; and the remaining cells (T5, T6, T8, and T9) consider some combination of the two program components. This example is purposefully expansive (nine cells) in order to highlight the potential that the factorial design has for answering interesting questions about the contribution of selected program components (and various intensities of them) to overall program impacts.

In order for a factorial design like this to be implemented in practice, program administrators and staff would need to be willing and able to ensure that each treatment group was provided specifically what it was randomized to. There are some applications for which this can be straightforward (e.g., mandatory versus voluntary meetings with a case manager, whether an individual is issued one of two versions of an assessment tool), and others for which it may be more complex (e.g., sequencing of certain program components). That said, the payoff in terms of knowledge gained is potentially substantial. Technical assistance to programs could help ensure faithful implementation and the evaluation’s overall success.

Before embarking on this—or any other evaluation design—it is important to conduct a power analysis to ascertain the needed sample sizes to precisely measure the effects that are relevant to policy and practice. As illustrated in Exhibit 3, the same sample size allocated across a factorial design increases the minimum detectable effect (MDE) size but also increases the number of

questions that can be addressed. Alternatively, a larger sample size is needed to be able to detect the same MDE.

- **Staged designs.** In a multi-stage design, participants initially randomized into a particular group are subsequently randomly assigned a second time into a second-stage treatment if they reach a certain trigger point. This design randomly assigns to treatment or control group at intake, subsequently observes eligibility for an aspect of the intervention in the treatment group, and then randomly assigns eligible treatment group members to gain access to the second stage intervention or not. Generally, a staged design uses initial experiences or outputs of the treatment group members—or even simply the passage of time—to determine later treatment options, to which individuals are randomized.

Example. Where internships are in high demand and require some initial program completion of early-program demonstrated success, they are suited for a staged evaluation design. Program participants must achieve some initial milestone to become internship ready, at which point they can be eligible for the internship lottery. The design permits estimating the effect of internships alone but also the overall effect of the program within which internships are a selective part.

- **Adaptive designs.** One particular type of staged design is that of a “sequential, multiple assignment, randomized trial” (SMART; e.g., Murphy, 2005). This occurs when a participant’s initial responsiveness to treatment determines later treatment options, considered adaptive treatments. In the case of career pathways programs, this is relevant because of the high level of customization that may take place: programs often aim to individualize services. This design provides the possibility of randomizing participants to alternative and responsive treatment options across the lifespan of a program. Adaptive designs can help assess the optimal sequencing of treatments or aspects of treatment as well as frequency or type of contact, for example.

Example. Maintaining students’ attendance in training can be challenging, especially with low-income populations, and various strategies can be used. For instance, a counselor can check in with each trainee on a regular basis. An adaptive experiment to improve attendance and guard against dropping out might involve randomizing trainees to receive contacts once every two weeks from a counselor (the treatment group) or not at all (the control group). After two months, attendance rates are reviewed. Treatment group members whose attendance was deemed insufficient (e.g., were absent more than once a week) would be re-randomized into two groups: one received contact every week from a counselor and one received it every two weeks. Those whose attendance was deemed sufficient also would be re-randomized into two groups: one continued to get contact every two weeks, and one for which contact was scaled back to once a month. After an additional two months, the impact of these adaptive interventions would be measured. This design permits testing the ideal frequency of contact, but the treatment may also involve sequencing of any kind of certain program experiences. To be administratively feasible, the implementation of an adaptive design could be mechanized, preventing staff from having to figure out who gets contacted when: simple reminders (set up to respond to the randomization results) can be set up to place contact actions on staff calendars, in the case of this frequency-of-contact example. More broadly, mechanized/computerized instruction on the appropriate service to provide a given client can help staff ensure adherence to the experiment’s protocols.

- **Cluster-randomized designs.** Finally, the designs defined above all use individuals as the unit of randomization; whereas cluster-randomized designs randomize on an aggregate unit—such as a school or a welfare office. To use this design to inform questions about the effectiveness of program components, *what* programs offer would need to be randomized at the program level. As such, programs would need to be both willing and able to permit a random process to determine what they offer. This seems impractical for fundamental issues of program design; but with respect to the “enhancements” or “options” that a program might offer, it could be a very powerful evaluation design for determining which program designs work best. When combined with individual-level randomization, a cluster-randomized design can be more powerful still.

Example. Research shows that working can negatively affect postsecondary persistence and completion by low-income adults. In response, programs could be randomly assigned to offer training stipends to their participants, perhaps in exchange for reduced work hours (e.g., not working more than 15-20 hours per week during training periods). Randomly assigning this component at the program level allows the treatment programs to embed the design into the program’s culture: focusing on career advancement and encouraging longer and/or multiple trainings that participants might otherwise see as infeasible in the absence of the stipend. Previous tests of financial incentives—such as wage subsidies—find that the marketing of the incentive is critical for its effectiveness (Martinson & Hamilton, 2011). The stipend could be conditional on meeting certain benchmarks of program participation and academic progress. Variations also can be tested: for example, some programs could offer the stipend only for time in education and training, whereas others could offer internships, too, with the stipend continuing during the internship.

It is important to understand the need of a control group under these designs. In the job training arena, it is not possible to have a strictly “no services” control group, in which people are precluded from accessing other services that are available in the community. Indeed, the “counterfactual” represents whatever else is available and posits treatment effects relative to those conditions. As such, the “control” conditions that are used as this benchmark can be thought of as a “business as usual” or *status quo* control group. This type of group is needed to determine whether a policy alternative improves on current policy options. However, this type of control group—where eligible individuals are turned away in order to learn about the relative effectiveness of the program in the current environment—is not strictly needed to ascertain whether a given program component improves on the standard program. If the commitment exists to deliver an intervention (no matter the research results), then testing an intervention against a control group may not be warranted. Instead, the evaluation can focus on comparison of alternative configurations of the intervention or vary certain aspects of the program in an effort to detect the effects of specific components.

To plan for a given new evaluation to assess the effectiveness of selected components of career pathways programs, DOL will need first to identify which components should be the focus of evaluation. With that, the specific component (or components) can be combined with the most appropriate evaluation design. With decisions about which components should be the focus of study, future planning can ascertain the best-fitting evaluation design and careful attention should be given to associated data collection needs.

### **Subgroup Impacts**

As with any of the designs discussed above, future studies planned for informing research questions about subgroup impacts must be designed with sufficient sample to permit detecting effects for particular subgroups of interest. Any such study should involve a preparatory power analysis to anticipate the sample size needed and the anticipated practice- or policy-relevant effect size. This may involve oversampling of priority subgroups to ensure sufficient power to detect effects, should they exist.

### **Sample Size Implications**

The prior discussion has identified several evaluation design variants aimed at answering certain questions beyond the simple question of whether a program works. Each of these has greater sample size demands than a basic treatment versus control experimental design. Any research project should start with its own power analysis to ensure appropriate sample size planning; and this subsection does not substitute for that careful work. It intends only to provide a general sense of the sample size needs of the alternative designs discussed above.

**Exhibit 3. Illustrative Minimum Detectable Effect Size Implications of Alternative Evaluation Designs**

Design/Test	Research Question	Treatment	Control	Total Sample in Analysis	MDE	Sample to Attain 3.9 MDE
<b>Basic Two-Armed</b>						
Treatment vs. Control	What is the overall impact of treatment versus control?	1,000	1,000	2,000	3.9 p.p.	2,000
50% Subgroup	Do impacts vary across the subgroup?	500 Subgroup 1 500 Subgroup 2	500 Subgroup 1 500 Subgroup 2	2,000	7.8 p.p.	4,000
<b>Three-Armed</b>						
Pooled Treatment vs. Control	What is the overall impact of either treatment versus control?	T1: 667 T2: 667	666	2,000	4.2 p.p.	2,753
T1 vs. T2/T1 vs. Control/T2 vs. Control	What is the impact of one arm versus any other arm?	T1: 667 T2: 667	666	1,334	4.8 p.p.	3,000
<b>Factorial</b>						
Any pairwise comparison	What is the impact of any treatment compared to any other treatment/control?	T1: 500 T2: 500 T3: 500	500	1,000	5.5 p.p.	4,000
Pooled Treatment vs. Control	What is the impact of any treatment (pooled) versus control?	T1: 500 T2: 500 T3: 500	500	2,000	4.5 p.p.	2,664
T1 and T2 vs. T3 and C	What is the impact of a specific component of treatment that is present in T1 and T2 but not in T3 and control?	T1: 500 T2: 500 T3: 500	500	2,000	3.9 p.p.	2,000
<b>Staged or Adaptive</b>						
T1 vs. Control	What is the impact of receiving T1 versus control?	T1: 1,000 → T2: 500 eligible Stage 2: 250 T; 250 C	1,000	2,000	3.9 p.p.	2,000
T1 & T2 vs. T1 only	For those who hit the trigger, what is the impact of receiving T1 and T2 versus T1 only?	T1: 1,000 → T2: 500 eligible Stage 2: 250 T; 250 C	1,000	1,000	6.4 p.p.	2,664
T1 & T2 vs. Control	For those who hit the trigger, what is the impact of receiving T1 and T2 versus control?	T1: 1,000 → T2: 500 eligible Stage 2: 250 T; 250 C	1,000	1,250	6.2 p.p.	3,125
<b>Cluster-Randomized<sup>a</sup></b>						
Treatment vs. Control	What is the overall impact of treatment versus control?	1,000 individuals in 30 sites	1,000 individuals in 30 sites	2,000	5.5 p.p.	2,000

Notes:

p.p. refers to percentage points.

All calculations make the following assumptions: a one-tailed alpha-level criterion of  $p < 0.10$  to statistical significance with 80 percent power; covariates explain 20 percent of variation in the outcomes. The binary outcome measure where the control group mean is 69 percent (which Peck et al. (2018) observe for the outcome of employment). The basic two-armed design assumes one-to-one randomization of 2,000 individuals. The three-armed design assumes one-to-one-to-one randomization of 2,000 individuals. Factorial design assumes 2,000 individuals randomly assigned to one of three treatment conditions or a control condition. Staged or adaptive design assumes one-to-one randomization of 2,000 individuals; then, within the treatment group, one-to-one randomization of the 500 individuals who flip the trigger to treatment 2. Cluster randomized design assumes 2,000 individuals evenly distributed across 60 sites with one-to-one randomization of sites.

<sup>a</sup> This calculation requires an assumed  $R^2$  at the individual and cluster levels and an assumed intra-class correlation coefficient (ICC). We assume 20 percent at both levels and an ICC of 0.02 (Nisar, Klerman, & Juras, 2013).



Exhibit 3 demonstrates the tradeoffs in statistical power across different design types and across different tests within the same design. For each design we assume a sample of the same size (2,000 individuals); we also apply the same assumptions to the parameters that underlie each power calculation. This ensures the variation in the MDEs is due only to the different designs and tests. Note that our assumption of equal chances of assignment to the various treatment and control conditions maximizes power under each design—an unbalanced assignment of individuals to treatment and control will reduce the power which essentially means that the minimum detectable effect size will increase. In general, the larger the study and subgroup sample sizes, the better able one is to detect a smaller effect or differential effect.

The basic two-armed randomization, with a sample of 2,000 cases (under the assumptions detailed above) can detect the smallest effect of 3.9 percentage points in employment levels (or other comparable binary outcome, for example). We also include a power calculation for a subgroup analysis within the basic two-armed design, in which the subgroup comprises half of the sample. In this case the MDE is twice as large as the overall analysis—7.8 percentage points versus 3.9. Although they are not included in the table, similar calculations for the other designs would also conclude that the MDE for the subgroup analysis is correspondingly larger than the MDE for the overall analysis.

Regardless of the particular test, a three-armed design has larger MDEs than the two-armed design. Put differently, three-armed randomization requires a larger sample than two-armed randomization to detect impacts of the same size. The same is true for the remaining designs—all have worse power than two-way random assignment (other than the tests within designs that are identical to two-way random assignment, such as T1 and T2 versus T3 and C in the factorial design). That said, these designs also offer the possibility to answer additional questions. In order to answer these additional questions, one must sacrifice the magnitude of the effect size *or* must accumulate a larger sample size. The far right column of Exhibit 3 shows how much more sample is needed in order to achieve the MDE of 3.9 percentage points, which is the smallest that we observe across the other designs, given our assumptions.

### 2.2.5 Quasi-experimental Designs and Analytic Approaches

Evaluations with a quasi-experimental design (QED), while not as rigorous as experimental evaluations, can play a role in developing knowledge about career pathways strategies.<sup>15</sup> QEDs do not randomize, as is done in experimental designs. Rather, QEDs attempt, using a variety of methods, to as closely approximate experimental conditions as possible, given that randomization was not used. These evaluations can cost less, often relying on existing administrative data (as opposed to tracking individuals using surveys) to determine programmatic impacts. They also can be used specifically for addressing the priority research questions stated here, considering long-term impacts, component impacts, and subgroup impacts. QEDs can be used when random assignment is infeasible or impractical, as occurs in some circumstances. In some cases, quasi-experimental evaluations can have greater external validity than experiments, as well. Common QED designs that would be relevant for

---

<sup>15</sup> In this report we make a distinction between “design” and “analysis.” This distinction corresponds to what those in the field of economics, for example, would refer to as “identification strategies” and “estimators.” We use the less technical evaluation language to discuss evaluation designs that can be put in place and provide the structure for identifying a program or policy impact of interest. Then, it is each design’s associated analytic approach that determines the details for how that impact is technically estimated.

this kind of research include reflexive and comparison group designs along with, more narrowly regression discontinuity (RD) designs. A “reflexive” design compares a treated group to itself at varying time points, such as with pre-post or simple (one group) interrupted time series designs. Comparison group designs identify a second group of observations, which again serve to approximate a counterfactual, or what would have happened in the absence of treatment. Regardless of the specific design (and we elaborate on several, below), quasi-experimental evaluation designs create a structure where a counterfactual is estimated.

The reason for approximating a counterfactual is to overcome threats to internal validity, which pose rival explanations for observed changes in outcomes. The most common “threats” in the job training evaluation world include selection bias (that those who access treatment differ from those who do not), maturation (that people mature/evolve in response to non-intervention stimuli), historical context (such as economic, political or social changes/trends), and regression-to-the-mean (that people at an extreme point—such as unemployment—are likely to move toward the mean, away from that extreme, naturally anyway).

With a chosen evaluation design, each then requires its own associated analysis to estimate impacts. After discussing the designs conceptually, we discuss the corresponding analytic strategies.

- ***Reflexive designs.*** A one group pretest-posttest design and simple interrupted time series design use a single group as a means to estimate a counterfactual condition: the pretest observation proxies for a counterfactual in the one group pretest-posttest design, and a pre-intervention trend intends to capture the counterfactual in a simple interrupted time series design. Reflexive designs are weak and generally do not account sufficiently for alternative explanations of program impacts. As such, we do not recommend their use in the job training arena.
- ***Comparison group designs.*** Adding a comparison group improves an evaluation’s ability to rule out rival explanations for observed impacts (what evaluators call “threats to internal validity”). Two main comparison group designs exist that could be used for evaluating career pathways strategies: a pre- and post-test comparison group design or a comparative interrupted time series. The *pre-post comparison group design* uses the pre-post difference in a comparison group to represent the counterfactual. Netting that difference from the pre-post difference for a treated group produces that design’s estimate of impact. Having additional baseline information is necessary to ascertain the groups’ balance, or lack thereof, so that analytic results can be interpreted with those group differences in mind. An improvement on the pre-post comparison group design is the *comparative interrupted time series design*, which involves having additional points of pre-intervention data to identify a trend. This baseline trend is what represents that design’s “counterfactual,” or what would have happened in the absence of the intervention.

The main challenge in using either of these designs is identifying a comparison group that is alike in as many ways possible to the group exposed to treatment. Recent research highlights the importance of using locally matched comparison groups; drawing a comparison group sample from national data will not usually yield a satisfying match (e.g., Jaciw, 2016).

Common practice in implementing a comparison group design is to use propensity score methods (PSM) (including matching or weighting) to simulate experimental conditions (e.g., Dehejia & Wahba, 2002). PSM involves selecting from the potential treatment and comparison group pools those cases that are similar on propensity scores, where “similarity” is ascertained by making one

of any number of choices. Propensity scores are themselves estimated using a model (typically a logistic or similar model) that estimates the “propensity” to fall in the treatment group based on a set of available characteristics of the subjects in question.

There are also alternatives to PSM for matching cases. There are methods that do not use propensity scores, but attempt to ensure, more directly, that the treated and comparison cases are similar on observed characteristics; that is, are balanced (King & Nielsen, 2016). There are many decisions to be made in executing PSM research, and to detail them all is beyond the scope of this report. Further, recent advances in this line of research application should be consulted before moving forward with a specific PSM strategy (e.g., King & Nielsen, 2016). Should DOL decide to use propensity score methods, an assessment of the various right choices for the research at hand should take place.

Example. Of the quasi-experimental studies listed in A.1, two used propensity score methods: Accelerating Opportunity (Anderson, Kuehn, Eyster, Barnow, & Lerman, 2017) and Integrated Basic Education and Skills Training (I-BEST; Zeidenberg, Cho, & Jenkins, 2010). The I-BEST evaluation, for example, looked at the relationship between participation in the program, which involved contextualized basic skills and vocational instruction for a population of students with low skills, and educational and labor market outcomes. Propensity score matching was used to match groups of students participating in the program with students who did not but were otherwise similar on a range of demographic characteristics and on income and test scores. The differences in the mean outcomes of the treated group and the matched untreated group were used to estimate the impact of the program, for each of the impacts considered: college credits earned and earnings. For robustness, the results of the estimate from the propensity-score-based models were compared to estimates from OLS regression and from a difference-and-differences model (see below).

- ***Regression discontinuity designs.*** Among quasi-experimental approaches, the RD design is the strongest for causal inference. In this design, selection to treatment is determined by a threshold or cut-point on a pre-treatment quantitative measure, the “assignment variable.” Study participants on one side of the threshold are assigned to a treatment, and those on the other side are not. To eliminate any bias introduced by selection, the assignment variable must be concrete and non-manipulable, meaning it cannot be altered directly by potential participants or staff, or based on preferences or staff discretion. At values of the assignment variable close to the threshold, the RD design is akin to an experiment with regard to internal validity. That is, those participants who fall just below and just above the threshold are likely similar in all ways, both measurable and unmeasurable, resulting in an unbiased estimate of the treatment impact at that point. Under some conditions and assumptions, the impact at the cut-point can reasonably be generalized to other values of the assignment variable.

Example. This design would be fitting to analyses of the effectiveness of career pathways programs in situations where program eligibility and access were determined using an assignment variable or variables. Achieving a given score on an entry exam, for example, is one possible means for programs to admit participants. A multi-faceted achievement score could take into consideration a wider range of skills, knowledge, and motivation (“fit”) for a program. To take advantage of this option for future research, screening assessments could be developed to yield a

cut-point score that would permit using the RD design to evaluate the effectiveness of the program.

We turn next to the analytic methods that stem from these designs: difference-in-differences models and individual fixed effects models.

- ***Difference-in-differences models.*** A difference-in-differences model is used to analyze the data from a comparison group design: examining the pre-post difference between the treatment and comparison group to yield an estimate of the treatment's impact. In addition to using this approach for analyzing individual-level data, a difference-in-differences model is also useful when there is a policy change at a specific point in time that does not occur everywhere. For instance, in addition to the PSM analysis noted above, this difference-in-differences model was used in the I-BEST study to estimate the impact at schools that had rolled out I-BEST versus ones that had not, thereby taking advantage of a multiyear statewide policy change (Zeidenberg et al., 2010). Cross-state analyses of changes in policy variation are relevant here, as well. Such models rely most importantly on the assumption that the trends would have been parallel in the treated and untreated units in the absence of the treatment. This assumption is stated concretely in the following example.

Example. Card and Krueger's (1993) well known study of the impact of the minimum wage increase on employment in New Jersey used a difference-in-differences method. They surveyed 410 fast-food restaurants in New Jersey and eastern Pennsylvania before and after the 1992 increase in New Jersey from \$4.25 to \$5.25. They assumed that the employment trends at the two sets of establishments would be parallel and looked at whether the trend at the New Jersey set was disrupted downward, that is, there was less employment than one would expect with parallel trends in both states. It was not, so they concluded that the minimum wage change did not affect unemployment. (In the career pathways context, the treatment could be the creation of a career pathways program in one state and the effects on employment in a set of employers in that state compared to a neighboring state that did not have such a program).

- ***Individual fixed effects models.*** These models are closely related to difference-in-differences models. In individual fixed effects models, there may be many units that are being compared, each of which has its own level and trend, whereas in difference-in-differences models, there are typically just two sets of units, each with its own mean and trend, as in the example of the two sets of fast food restaurants in the two states above. For both kinds of models, there are at least two time points.

Example. Consider 1,000 workers who participated in a career pathways program, where three predominant components were in operation; for instance, occupational training, career counseling services, and job placement services. Data that are available include a measure of wages in an initial period, in which each worker may have received some of the services (possibly none), and a measure of wages in a second period, in which each worker again may have received some of the services, but perhaps a different set. We will therefore have 2,000 wage measurements (each worker has two observations), which can be analyzed as a function of (1) the presence of three explanatory variables (the indicators for receipt of each of the services in the period prior to the wage measure) and of (2) an indicator for each worker (except one, to avoid perfect collinearity, or not having enough variation in the data) in the sample (999 indicator variables). These latter

variables are the “individual fixed effects.” Under certain assumptions, this analysis controls for the time-invariant individual-specific characteristics of sample members, and thereby creates estimates of the effects of each of the components of the career pathways programs on the wages. Note that the number of periods in the model is arbitrary; one could have three or more periods instead of two, but at least two are needed because the analysis relies on having longitudinal data.

For this and other identification strategies, outcome variables other than wages are possible, as well: an indicator variable that signals a transition to a target occupation or set of occupations, or another marker of career advancement, also can be analyzed as a function of the program characteristics experienced by individuals in the sample. This would allow a researcher to determine the effects of selected program components (or a general program enrollment) on career advancement. Again, both the program components and the outcome need to be observed for a single worker at two or more time points. Individual fixed-effects models do not need to assume a parallel trend between individuals. For instance, more educated workers might see more rapid increases in their wages. The individual fixed effects in theory control for both the unobservable and observable characteristics of each individual. However, the model does not control for characteristics of each individual that change over the time period. So, for instance, if a worker got more education over the time period, that would not be accounted for by the fixed effect (although it could be incorporated in the model elsewhere.)

In sum, analyses using regression discontinuity designs, comparison designs (including those employing propensity score matching), difference-in-differences models, or individual fixed effects models are appropriate when one has the right kind of data from a program or set of programs, and when the assumptions that are made by the model selected can be tolerated. Although these types of analyses do not allow a researcher to make fully causal conclusions, they are viable approaches for understanding the relationship between individuals’ program experiences and their outcomes. When experimental evaluation designs are infeasible, quasi-experimental evaluations and the associated analytic strategies are an option for examining issues related to programs’ long-term effects, component effects, and, data permitting, subgroup effects.

### 2.3 Data Sources

This section summarizes data sources that should be considered for studies of impacts of career pathways approaches.

#### 2.3.1 Existing Impact Evaluations

In planning future evaluations on career pathways approaches, DOL should consider the existing data sources that would allow researchers to effectively analyze them. Previous evaluations have used a variety of data sources which may also be useful in future work on career pathways. Among these data sources, the ones used for impact evaluations using randomized experiments appear to be the most useful for extended career pathways evaluations in the future, because they have (or had) regularly implemented data collection procedures for outcomes data.

In Exhibit A.1, we summarize completed and ongoing impact evaluations relevant to career pathways strategies as designated by the type of design—either experimentally designed evaluations or quasi-experimentally designed evaluations. Because DOL is interested in understanding how the workforce field itself defines and implements career pathways initiatives, our work in the knowledge

development phase cast a broad net in deciding which studies to include. The *Research and Evaluation Synthesis* report included any study that included initiatives that described themselves as involving career pathways, as well as others that had at least some elements of the pathways approach, whether or not they met any particular definition of the model. The report also included several training efforts that were not explicitly career pathways approaches—in some cases predating the concept—where there were major studies of initiatives representing core elements of that model, such as sector partnerships or integrated basic education and training.

Although evaluations of program components that became part of the career pathways approach date to the 1980s, there has been a recent wave of research that has evaluated larger-scale programs, including ones that explicitly adopt a career pathways approach, usually implemented at the federal level, with larger sample sizes. Many of these evaluations have implemented experimental evaluations, which offer strong causal evidence on impacts of these career pathways programs. Many of the completed impact evaluations had follow-up periods of four years or less, with the majority three years or less. In large part, this was because tracking individuals after randomization was costly and limited by the length of an evaluation contract. The ongoing impact evaluations generally have longer follow-up periods, and so ultimately will offer more insights into long-term impacts than the completed ones. However, most ongoing studies have published only short-term impact findings so far. As noted earlier, just two studies to date have reported long-term impact findings (five years or longer).

The Pathways for Advancing Careers and Education (PACE) evaluation provides an example of a current career pathways study using an experimental evaluation design that includes relatively long-term follow-up. PACE is evaluating a diverse group of nine career pathways programs in a range of industries across the United States with different settings and target populations. The PACE impact evaluation consists of an experimentally designed impact evaluation at each site, with randomization occurring at the individual level: students were randomly selected either to participate in a career pathways program or not. The PACE evaluation has one of the larger sample sizes among the studies in Exhibit A.1; the evaluation of the Pathways to Healthcare Program at Pima Community College in Arizona, for example, has a single-site sample size of more than 1,000 students. The PACE data include administrative records on study participants (e.g., community college course enrollment, credential completion, employment and earnings data), as well as a follow-up survey that measures varied outcomes, including service receipt and job quality and progression. The early impact reports measure outcomes at approximately 18 months after random assignment, and a follow-on study is measuring outcomes at 36 and 72 months.<sup>16</sup>

This type of evaluation would be an ideal target for even longer-term analyses on outcomes following career pathways advancement, and it would help address the *Long-term Impacts* research question on advancement in career pathways. It would be necessary to match existing administrative data sets with the records that have been collected through an evaluation such as PACE, where information on administrative wage records for a longer period of time can be collected. This would help determine, for example, whether individuals who were randomly selected to participate in these career pathways programs experienced higher earnings, wages, and other outcomes of interest potentially five or more years after program entry.

---

<sup>16</sup> PACE reports published to date are available at <http://www.career-pathways.org/>.

The ability to measure long-term outcomes is important when considering quasi-experimental designs to evaluate career pathways programs. Exhibit A.1 also details previous and current examples of quasi-experimental designs, which do not delineate treatment and control conditions through randomizations. Although there is a lesser ability to make causal claims from results because of selection bias, the sample sizes are often larger, because it is possible to create a comparison group using administrative data. With more administrative data on individuals becoming more readily available over the last decade or more, and considering the reduced cost for these types of evaluations, QEDs can provide answers to questions retrospectively on impact measures without the time and effort necessary to set up an experiment prospectively.

The example of the I-BEST QED evaluation, discussed earlier, highlights how early promising results from a QED were built on with a longer-term experimental evaluation to more accurately capture successful long-term outcomes. One drawback of the evaluation's propensity score matching and difference-in-differences analyses of statewide administrative data was the limited timeframe of the outcome measures, which was only two years from the start of college enrollment. The I-BEST program is now undergoing another round of evaluation in PACE, where impacts are being measured using an experimental design and a longer follow-up period to measure outcomes (Glosser et al., 2014).

In addition to linking to administrative data, DOL also can consider approaches in linking existing career pathways program evaluation data to newly developed surveys or qualitative means of collecting data. For instance, it might be worthwhile to follow up with individuals who were a part of selected evaluations with new surveys on employment outcomes or other outcomes, many years removed from the time of their randomization. These types of data collection efforts also would be helpful in determining whether individuals' long-term employment outcomes change over time, and whether these outcomes differ across the treatment and control groups.

In addition to linking data, DOL could consider pooling data from existing evaluations. One method of doing such pooling is to take data from across studies where both the individual-level and program- or site-level data are available, and analyze how program variation associates with variation in individual impacts. Such an analysis can help determine the contribution to program impact of individual program components and implementation practices. In existing evaluations that are implemented across various institutions, it can be difficult to determine the specific impact of a program component without pooling data—in order to have enough program- or site-level variation to analyze. To appropriately pool these types of data, consistent program-level measures and outcome measures must exist across the evaluations. As noted in Subsection 2.2.1 on descriptive analyses, one option would be for new research to collect retrospectively these program-level measures for use in either pooled individual-level analyses or meta-analyses.

### **2.3.2 Administrative Data**

Beyond existing impact evaluations on career pathways strategies and linking to their data sets, another option is the use of primary administrative data, mostly for exploring the *Component Impacts* and *Subgroup Impacts* questions. Administrative data can be used with an appropriate causal research design—such as a randomized controlled trial, leveraging of existing lotteries, or a difference-in-differences approach—to answer questions about the impacts of programs on participants' outcomes, but cannot reliably answer such questions in the absence of a strong design (Klerman & Rolston, 2015). Large-scale administrative data sets can also help determine which program components are

most likely to lead to positive outcomes, and which subgroups might gain the most from career pathways programs.

In order to link data between different data sets, whether or not the data are administrative or are drawn from surveys, one typically makes use of identifiers of various sorts. These can be numeric identifiers such as social security numbers and names. This information is sensitive and needs to be safeguarded under guidance from one or more Institutional Review Boards (IRBs). Also useful in matching records are such data as date of birth and gender. Matching can be exact or inexact; the latter is often useful when there are errors in one or more data sources. Matching can also be one-to-one, one-to-many, or many-to-many, depending on how many records a particular person has in a particular data source. For instance, in an earnings database, each individual may have one record per quarter.

The fundamental source for these types of analyses is administrative, which is individual-level data that generally have good coverage (meaning data are available for the vast majority of the study sample) and can have minimal missing values. Several potential sources are available to not only track outcomes, but to incorporate detailed demographic information, as well. The National Directory of New Hires increasingly is used among researchers to track individuals' wage and employment histories over longer periods of time. NDNH also has the advantage of tracking wages and employment histories across state borders (which UI data is unable to do, because it must be requested from individual states, although requests from multiple states are possible).

In addition to NDNH and UI, DOL could explore the use of Internal Revenue Service (IRS) data to track individuals' earnings, other income, and employment histories. The IRS has both earned and unearned income data on individuals. IRS data is increasingly being used in research, although the process of obtaining the data is still long and involved and can be costly.

Lastly, DOL might explore options in ECE or postsecondary data, to provide a more complete picture of individuals' educational histories and trajectories, both before and after their enrollment in a career pathways program. Options include the T.E.A.C.H. data, which includes information on those who pursue careers in ECE and their education and employment outcomes. Additionally, the National Student Clearinghouse (NSC) is a comprehensive data set that provides information on enrollment and certificate and degree attainment on almost all postsecondary students at public and not-for-profit private universities and colleges in the United States (participation by for-profit postsecondary institutions is more limited). It is important to note, however, that NSC coverage of non-credit enrollments and credentials is less complete than credit-bearing ones. These data may be able to help researchers understand the extent to which individuals are completing multiple stackable credentials in a career pathway, and determine whether students are progressing through career pathways programs by completing related credentials, at least in the kinds of institutions that the NSC covers.

### **2.3.3 Survey Data**

DOL could consider collecting new survey data on selected outcomes and/or specific populations of interest. In learning more about individuals who enter career pathways and attempt to advance through them, it is important to collect follow-up information on their outcomes, particularly in aspects that are not regularly collected through administrative data sources. The knowledge and



understanding of their potential career and employment options,<sup>17</sup> the quality of the jobs participants obtain, the extent to which they continue on in education and training, and the extent to which they can balance personal and professional responsibilities are some of the important pieces of information that could shed light on interpreting the impacts of career pathways programs and help generate theories about ways to improve their effectiveness. They could also be used to collect information on underrepresented groups, where less information is currently available (see Section 2.4 below). A key issue will be size of the survey and whether it can be conducted to be representative of the existing sample.

As noted above, possible new survey data might include long-term follow-ups with participants who were part of a completed career pathways study. New data collection presumably would involve some type of linkage to that evaluation, with the new surveys asking about their employment outcomes and whether their perspectives on jobs and careers had changed after several years. New surveys also can include the perspectives of employers, or the American Job Centers staff who are employed to help workers seeking employment or additional training.

Existing survey data also can be useful. Valuable insights can come from cross-sectional data sets that contain occupational, wage, and educational information. One such survey is the Bureau of Labor Statistics (BLS) Occupational Employment Statistics (OES).<sup>18</sup> The relationship among occupations, education, and wages found in such data sets provides the context required for the planning and implementation of career pathways programs; potential and current program participants also use this important information to inform their own career planning and choices.

### 2.4 Practical Considerations and Tradeoffs

The impact questions for future career pathways research are important. Indeed, future research has the potential to illuminate the long-term effects of career pathways strategies, as well as provide insights about which specific components of career pathways drive impacts and about subgroup effects. Adding to existing research and embarking on new research need not be mutually exclusive endeavors; but it is important to prioritize which program components and which subgroups are of greatest interest to DOL and to the field. That information is essential in order to design plans that build on existing evaluations (by reanalyzing existing data or conducting additional follow-up) or plans for new evaluations. With this in mind, there are some tradeoffs for DOL to consider.

#### Sample Size and Associated Detectable Effect Size

As discussed, evaluating the impacts of many program components and component-subgroup combinations requires sample sizes of sufficient power to detect component-specific and subgroup-specific impact magnitudes. Sample sizes have direct implications for the cost and practicality of conducting new evaluations. If DOL chooses to plan new evaluations, we would encourage a more specific power analysis than what we present here to identify the optimal sample sizes for the specific research. Although larger studies are simply more expensive than smaller ones, there are certain set-up costs regardless of sample size; once these are sunk, the incremental cost of additional respondents

---

<sup>17</sup> Knowledge and understanding of career and employment options is hypothesized to improve employment outcomes through counseling provided in career pathway programs (see Fein 2012).

<sup>18</sup> Such surveys typically are of employees and may not fully represent the self-employed. The use of such surveys in industries such as early childhood education where many workers are self-employed may therefore be limited.

is marginal. Various factors affect those set-up costs, however; a new evaluation that requires a large individual-level sample also might need to take place in multiple locations. Planning for a study's sample size—and the implications of that size for costs and logistics—should be informed by power calculations. Moreover, an evaluation that has differential or additive effects (of various program components, for instance) can expect to have relatively smaller detectable effects than an evaluation that considers an overall treatment effect; and these smaller differential or additive effects require a larger sample to be detected.

### **Focus on Components**

DOL also must consider what will be necessary in order to evaluate the effect of program components of interest on career pathways programs. For new components, programs may need new funding to be able to integrate the components into their practice. Further, new program components need to be implemented in a way that permits their rigorous evaluation. The HPOG Impact Study's addition of three such components in 19 of its 42 programs illustrates that this is possible. To evaluate the effect of existing program components may not require additional funding, but instead will require programs to reconfigure how they provide participants access to the components, aligning with research needs while minimizing program disruption. These are not insurmountable challenges, but ones that DOL must consider and plan for to ensure the success of future research.

### **Consent Language**

Another category of practical considerations relates to informed consent of study participants. Research ethics requires that people enter into studies fully informed and voluntarily. As a result, research uses “informed consent” agreements, which advise participants of their rights, expectations and possible risks. Usually, these agreements detail what data will be used, how data will be used, and who has access to it, and typically limit it to a specified period of time. As a result, to re-analyze or build on existing evaluation data requires researchers to examine past informed consent language to ensure that further work and/or additional data collection is permissible. If not, the evaluation would need to secure new informed consent agreements, and the feasibility, cost, and time involved in doing so would need to be explored. At minimum, because of the promise of using existing data or collecting additional data that builds on previous evaluations as a cost-effective option for examining the correlates of program impacts, future research projects should use consent agreements that permit longer-term follow-up and add-on studies funded by different contracts.

### **Focus on Underrepresented Groups**

A final set of research challenges relate to focus on underrepresented groups. Groups that have not been part of career pathways research in the past also are not readily available as the subject of future research because they may continue to be underserved by career pathways programs. To address that in future research will require modifications within programs themselves. Changes in outreach strategies, changes to target participants, changes to performance management systems, and changes to make programs more accessible to a diverse participant pool are not simple to make. Existing programs will need to make them, however—and new programs will need to be designed accordingly—if future research is to result in lessons learned about how otherwise underrepresented groups experience career pathways programs.

Qualitative research methods, used in conjunction with some of the approaches outlined above, could be particularly useful for identifying and studying underrepresented groups. Such methods could focus on the barriers they face to participation, completion, and employment, and the program

components or configurations of components that may improve their career pathways program outcomes. These research approaches might include surveys of and interviews with underrepresented groups. The PACE evaluation, for example, has conducted a series of interviews with a small sample of treatment and control group members about such questions. While not representative of the entire sample, key findings include that though a majority of participants were still on the career pathway 18-24 months after enrollment (either working or training in the field), they no longer saw themselves as connected to the program and faced a number of challenges including tight finances, low-paying jobs, and difficulties balancing training and family responsibilities (Seefeldt, 2017). These findings provide important context for understanding the study's findings on impacts.

### 3. Workers' Career Trajectories in the Economy

This chapter considers descriptively workers' career trajectories. This is defined as the series of occupational steps that correspond to job opportunities, generally with higher wages, that exist in the labor market, regardless of specific programs or initiatives.<sup>19</sup> Research on career trajectories can be used to set the context and expectations for what might be achieved in career pathways programs based on the experience of workers in the labor market more generally (not just those in career pathways programs). Specifically, a deeper understanding of the extent to which and how workers advance in the labor market in particular sectors and occupational clusters could help policymakers and practitioners better target career pathways initiatives to those that offer the most potential for advancement. Such research could also inform design of career pathways programs by examining the steps typically followed by workers who succeed in advancing, as well as how common or rare progression through those steps is.

The motivation behind this research is the following: If in examining the labor market as a whole—or a sub-segment of it such as a large healthcare system—workers do not frequently make a transition between a particular pair of occupations, that raises the question as to whether it makes sense to include such a transition as a goal for a career pathways program. Conversely, if we observe a particular occupational transition frequently, then planners of career pathways programs may want to consider including it in a career pathways program, especially if the transition appears desirable (e.g., represents an advance in skills, wages, or both).

In other words, because the labor market as a whole is the context in which career pathways programs operate, it makes sense to look at the transitions in the labor market as a whole in order to understand this context. Such analysis also could help generate and refine research questions for career pathways evaluations.

#### 3.1 Introduction to Career Trajectories Research Questions

The research questions that this chapter considers are the following:

*In the absence of career pathways programs, to what extent, and how, do workers advance on their own through multiple, progressively higher levels of education and training, and associated jobs, within a pathway over time? To what extent is advancement more common for workers from specific settings or backgrounds? To what extent do wages increase for workers who progress through these pathways? What are the implications for designing and evaluating career pathways initiatives?*

In addition, policymakers and practitioners seek to understand how career trajectories might vary by business and industry sector and by occupational cluster. For example, Congress specifically highlighted its interest in the healthcare and ECE sectors in WIOA's language requiring DOL to conduct career pathways evaluation research. In response, this chapter also considers approaches for generating useful research on the ECE sector specifically (per DOL's request, given the scarcity of

<sup>19</sup> Occupational steps may involve changes to jobs with lower wages, or to jobs at the same wage, with other features that compensate for the lower earnings, such as more and better training and advancement opportunities, or a better work environment, commute, or hours, etc. However, given the focus in the career pathways field and in the WIOA legislation on low-wage workers, transitions that lead to higher wages and earnings are of the most interest for this report.

information on it relative to the healthcare sector). We also include healthcare examples in this chapter, however.

Career trajectories are composed of occupational transitions, and several factors influence the frequency of a particular transition between two occupations.<sup>20</sup> The first is the overall frequency of each occupation in the economy as a whole. Some occupations are common, in that there are many incumbents; examples include Retail Salesperson, Cashier, Food Preparation and Serving Worker, Office Clerk, and Registered Nurse (these were the top five occupations in BLS's Occupational Employment Statistics in May 2016).<sup>21</sup> OES also tracks many less common occupations. Some of those found least frequently are Household Cook, Fishing and Hunting Worker, and Radio Operator.

All things being equal, if an occupation is large, then it will attract more people into it because there will be more openings (even if it is not generating openings at as rapid a rate as a small occupation). At the same time, if an occupation is large, more people will be leaving it. In fact, openings and departures, rather than the size of the occupation itself, mostly determine the number of transitions. If an occupation is large but it is not growing, and people are not leaving it at a high rate due to retirements and quits, it may not offer many opportunities despite its size. BLS estimates openings for occupations as part of its employment projections program, estimating both employment growth (or decline) and turnover. However, BLS does not estimate what occupations the people filling the openings come from.<sup>22</sup>

The second factor that affects the frequency of a transition between two occupations is their similarity. Occupations can be similar in several ways. They can be similar in the skills and/or education they require, in wages, and in industry sector or type of establishment (e.g., hospital, construction site). Note that one of the main goals of career pathways programs is to facilitate transitions from lower-paying jobs to higher-paying jobs, typically within the same sector. The "lifelong learning" movement also drives workers to get more education to move to more skilled occupations and improve their status (Coleman, 2017). It is important to know the extent to which such transitions are taking place in the labor market as a whole in order to know the extent to which one can realistically expect them in the context of career pathways programs. Such analyses are also important for assessing the extent to which local successes in career pathways programs are scalable to the entire economy.

### 3.2 Research Approaches

This section discusses five research approaches to understanding workers' career trajectories:

- Using extant survey and administrative data to identify trajectories that occur commonly in the labor market,

<sup>20</sup> In some cases, there can be a career ladder within a single occupation in a particular employer's occupational definitions (e.g., "Scientist," "Senior Scientist,"), but we are focusing on transitions between occupations in this section.

<sup>21</sup> Information on the 10 most prevalent occupations is available here: <https://www.bls.gov/oes/2016/may/largest.htm>.

<sup>22</sup> It is very common for workers to move from one job to another within the same occupation, but this movement is not of much interest for career pathways programs, except to the (substantial) extent that the movement typically prevents others from entering that occupation from other occupations.

- Designing a new survey instrument to track workers using a sample that is representative of the labor market as a whole or a single industry sector,
- Using predictive analytics to determine possible career trajectories of workers,
- Conducting cross-state difference-in-differences analyses to determine the effects of various policies on transitions, and
- Conducting qualitative studies based on interviews with workers to understand individual workers' trajectories and choices.

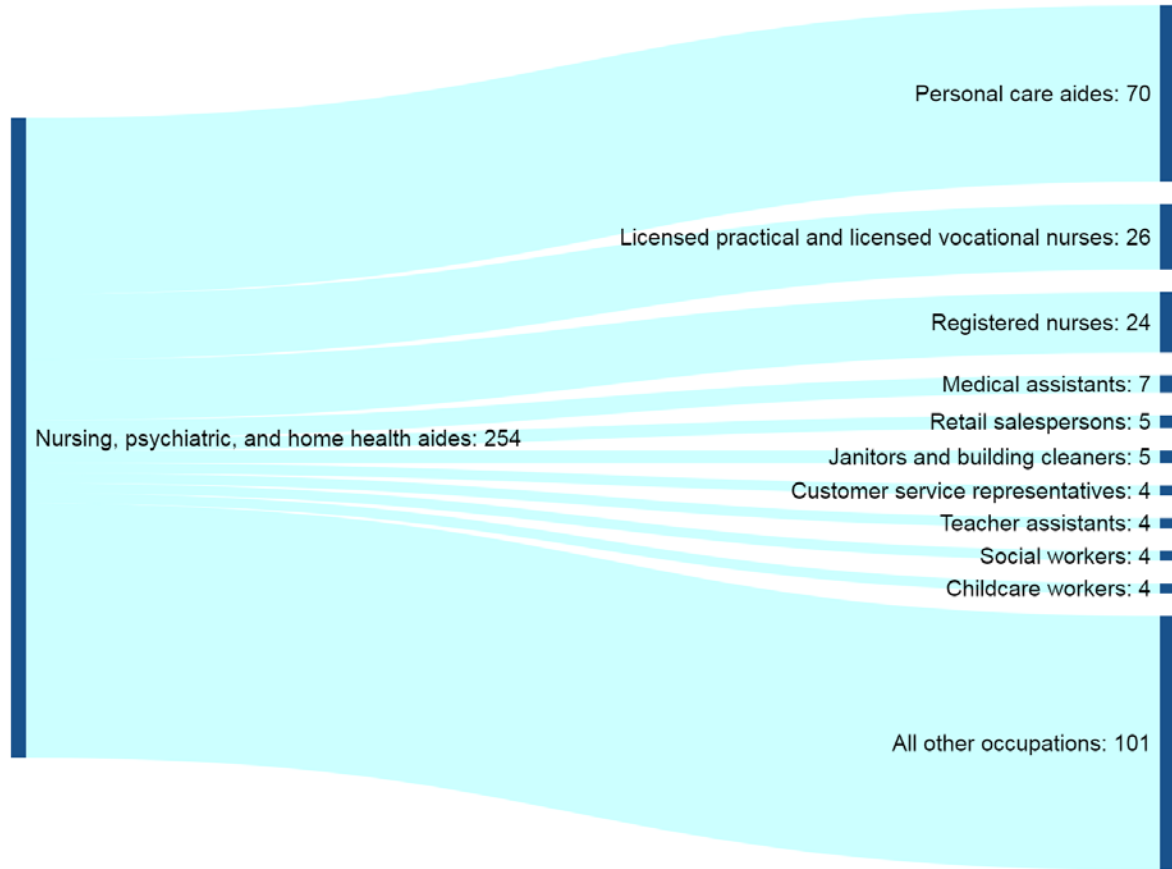
### 3.2.1 Using Extant Data to Identify Common Trajectories

Two types of extant data—survey and administrative—can be used to understand descriptively the occupational transitions that are occurring in the economy. Section 3.3 provides detail on sources of each of these types of data.

In using extant survey data, it is important that the data set be large, representative, and have a longitudinal component so that occupational transitions can be captured. Extant federal government surveys include the Current Population Survey (CPS from Census and BLS), Occupational Employment Statistics (from BLS), and the National Longitudinal Survey of Youth (NLSY from BLS). One problem is that unless an occupation is large, not many incumbents of that occupation will be present in the type of representative sample typically found in such surveys. This problem is exacerbated when the research focus is only a single industry. This can result in unreliable estimates when transition counts are low (Kambourov & Manovskii, 2013). Typically, extant surveys of this type do not focus on particular sectors of the economy, but the economy as a whole. The NLSY and CPS contain information on educational attainment and demographics that can be used to better understand who is undertaking these transitions.

For a previous DOL project, Abt used data on occupational transitions from the Current Population Survey. Looking at one occupation often included in career pathways programs—Nursing, Psychiatric, and Home Health Aide—we found that 254 incumbents transitioned from year to year into another occupation. The most common transition was to Personal Care Aide (70 cases), followed by Licensed Practical and Licensed Vocational Nurse (26 cases) and Registered Nurse (24 cases). These two transitions to nursing often would be included in a career pathways program. Exhibit 4 displays the findings about transitions in a Sankey diagram, a type often used to visualize flows.

**Exhibit 4. Example – Sankey Diagram of Occupational Transitions of Nursing, Psychiatric, and Home Health Aides**



Source: Analysis of Current Population Data (CPS).

Information such as this would be valuable to designers of career pathways programs in understanding what occupational transitions are most likely and therefore perhaps most feasible to incorporate into a program. For example, career pathways “roadmaps” often show Registered Nurse (RN) as the career step following Licensed Vocational Nurse (LVN). In a similar analysis of LVN incumbent transitions for the same DOL project, however, we found that workers more commonly moved to one of a variety of Technician jobs.

**3.2.2 Tracking Workers with a New Survey of a Representative Sample**

A new survey could track workers across their careers, examining their wages over time, their occupational trajectory, firms worked in, and education and training credentials earned. It also could collect demographic data. Typically, a new survey would be targeted at a particular sector, such as healthcare or ECE. The ideal survey would be longitudinal (covering as many years of a worker’s career as possible) with data collected at regular enough intervals to minimize issues with respondents’ recall. Such a survey would have a large sample size to allow for estimates from sectoral subsamples. As a result, such a survey would be quite costly. Though focusing on a particular sector may be less expensive, it would still be quite costly because it would need responses from sufficient numbers of workers in each occupation in that industry to gather reliable statistics on transitions and

also to represent the whole country and smaller geographies such as states. Relatively less common occupations might need to be oversampled.

Another possibility is adding occupational data to UI and LEHD data. This would be a new mandate to states and employers, and as such might be difficult to achieve, but access to such data would greatly increase the value of UI and LEHD data to policymakers and researchers.

Thus, there are tradeoffs between representativeness and cost. It may be more cost effective to conduct a survey in a single large firm, such as a large healthcare system, even if employment in that firm is not fully representative of its industry as a whole. It also might be necessary to oversample particular occupations, if possible; especially those occupations that might not be sufficiently represented in a simple random sample.

If cost concerns can be overcome, surveys can be a source of rich data. For instance, researchers could survey workers in detail about their education and skills (the latter using a typology of skills such as exists in O\*NET, DOL's preferred system for describing occupations). If the survey is longitudinal, and measures career trajectories in terms of occupations, industries, and wages, interesting studies can be made that relate the measured education and skills to the trajectories.

Custom surveys are probably more practical when they focus on occupations in particular industries or within particular regions. Lists of employers within a particular industry are useful for conducting such surveys. Dun and Bradstreet is one of the best sources of employer lists, and its lists can be searched geographically and by industry, since every employer has an associated industry (NAICS) code. Once researchers have made a list of employers, they can be surveyed or their permission sought to survey their employees. This can become logistically difficult and costly when many employers are involved, especially if many of them employ small numbers of workers. For instance, the ECE field has a mix of employers, from large national chains such as Bright Horizons to small local chains and stand-alone centers, home care, and nannies. It is much easier to survey just one or a few large employers than to try to survey a whole industry, but at the expense of representativeness, since a few employers are not representative of an entire industry.

### 3.2.3 Predictive Analytics to Determine Promising Career Trajectories

Researchers are increasingly interested in the use of predictive analytics and machine learning (Kelleher, Mac Namee, & D'Arcy, 2015) to improve the way programs function in a wide variety of policy domains, including in workforce development. *Predictive analytics* uses information about the past and present to predict the future. One common application of predictive analytics is in credit scoring, in which an individual's history of repaying debt is used to predict their credit worthiness. *Machine learning* is a field of computer science where computers learn from available data without being explicitly taught, but rather through examples. For instance, a machine may learn to recognize a cat by showing it two sets of pictures, one set that contains cats, and one set that does not, where the pictures are labeled as positive and negative examples (Clark, 2012). Machine learning and other statistical techniques are commonly used in predictive analytics.

One idea is to apply predictive analytics to available data, to predict the likely career trajectories of individual workers (or subgroups of workers) using their demographic characteristics and their wage and employment histories. For instance, one might predict the future earnings of a worker who is a Home Health Aide, based on current earnings, test scores, education, and age, as well as the probability of transitions into related occupations such as Medical Assistant, Licensed Practical Nurse



(LPN) and Registered Nurse. Such analyses can generate recommendations on the type of training those workers should receive from career pathways programs to optimize their wage or employment outcomes in the future. For instance, if it is highly probable that person can transition to licensed practical nurse, LPN training might be recommended. One must be cautious that the use of such analytics do not have the effect of reproducing existing demographic segregation in occupations. For instance, if men are not found in many health professions, one would not want to discourage a man from going into such a profession because the analytics “predict” he is not suited for it. Similarly, one would not want to discourage women and/or members of racial and ethnic minority groups from going into construction, or other fields in which they are under-represented (Berger, Black, & Smith, 2001).

In another example, a worker’s retrospective occupational and wage trajectories can be used to predict possible future job options for the worker and perhaps could be used in American Job Centers (AJCs), in educational contexts, and elsewhere. Education, skills, experience, and potentially other characteristics of a particular worker could be matched to similar workers in a data set and used to suggest possible next steps in their career pathway. These suggested next steps could be customized to a particular local labor market, as local conditions vary greatly. The process also could be used for “what if” modeling, so individual workers can see their likely future employment and wage trajectories based on what kinds of training they receive. Data from private vendors such as LinkedIn and PayScale also can reveal relationships between particular educational choices and career outcomes. Initiatives designing career pathways also can use these retrospective trajectories to determine the most promising sequences of education and training to invest in for the future.

### 3.2.4 Cross-State Difference-in-Differences Analysis

Different states have different policy environments for similar occupations. For instance, one state may have more stringent licensing requirements for a particular occupation than does another, and this can influence the flow of workers into or out of occupations. If two neighboring states have similar regulatory environments, and then one of them makes a policy change to a more or less stringent environment, a natural experiment<sup>23</sup> is set up, which can be analyzed using the difference-in-differences method described in Section 2.2.5 on quasi-experimental methods. Such research has the possibility of illuminating the extent to which policy change can affect the magnitude of the flows of workers along particular occupational transitions.

For example, there has been pressure in recent years to increase the formal educational requirements for Registered Nurses. Formerly, an RN could have either a two-year or a four-year degree; more and more, however, healthcare employers are requiring a four-year degree. If one of the states began requiring that RNs have a bachelor’s degree to be licensed, then the change would affect flows into the occupation, wages in the occupation, and the supply of RNs available to work in that state. A difference-in-differences model comparing states with such a rule to states without it could estimate these impacts.<sup>24</sup>

---

<sup>23</sup> The term “natural experiment” is typically used to describe such situations, but it is important to note that such “natural experiments” are not true scientific experiments, because the units of analysis are not randomly assigned to the treatment and control conditions. Therefore selection bias is still present.

<sup>24</sup> One might want to omit states adjacent to the treated states in such an analysis in order to avoid concerns about spillovers (Black, McKinnish, & Sanders, 2005).

### 3.2.5 Conducting Qualitative Studies to Understand Workers' Trajectories

Qualitative studies could also be useful for understanding how workers progress through career trajectories. Structured interviews with workers in selected industries could inquire about the professional and education and training experiences of workers who had advanced to their desired occupations, as well as workers who were hoping to take such steps. The focus of such a study would be on understanding workers' aspirations within an occupational field, the steps they took or plan to take toward advancement, barriers they encounter, and the services and supports they need to successfully advance. The PACE treatment and control group interviews noted earlier are an example of the rich information about barriers and facilitators that can be gathered in this way (Seefeldt, 2017).

Such a study could be useful in understanding trajectories in fields in which advancement appears limited. For example, within the early care and education field, Cheng et al. (2018) noted that many child care workers who obtain additional education leave early care settings for the K-12 system. A qualitative study could help researchers to understand whether workers would prefer to stay in non-school early care settings and if there are barriers to doing so beyond lower wages. Qualitative studies could also be useful for understanding whether trajectories look different for different subgroups of workers, and could illuminate barriers for subgroups that tend to be less successful at advancing in a given occupational field.

Given that such a qualitative study would likely involve a relatively small number of workers and would be limited in terms of representativeness and generalizability, it may be done in conjunction with one of the other approaches outlined above.

## 3.3 Data Sources

This section reviews seven survey and administrative data sources relevant to research on career trajectories: the Current Population Survey, Panel Study of Income Dynamics, American Community Survey, Occupational Employment Statistics, data collected as part of the Workforce Innovation and Opportunity Act, the National Longitudinal Survey of Youth, and data from private companies. This section also describes structured interviews that could be part of a qualitative study aimed at understanding individual workers' trajectories.

### 3.3.1 Current Population Survey

The Current Population Survey is a joint effort of the BLS and the Census Bureau and represents the main source of timely information about workers' wages, occupations, and employment status. It is a monthly survey of about 60,000 households designed to be representative of the country as a whole. Households are rotated in and out of the survey each month. A given household is in the survey initially for four consecutive months, then out for eight months, and then in for four more consecutive months, before leaving permanently. Thus, each household is tracked in two consecutive years, allowing for a possible job transition.

There are various supplements to the CPS. Perhaps of greatest interest is the Job Tenure and Occupational Mobility Supplement, typically conducted in January or February, which asks workers about their occupation a year ago. Most workers do not change jobs over the course of a year, but some do, and occupational transitions can therefore be mapped with this data. Anonymized CPS

microdata are available from BLS; versions are also available from some third parties, such as the National Bureau of Economic Research.

### 3.3.2 Panel Study of Income Dynamics

The University of Michigan directs the Panel Study of Income Dynamics (PSID). Begun in 1968, it contains a nationally representative sample of over 18,000 individuals living in 5,000 families. Information on these individuals and their descendants has been collected continuously, including data covering employment, income, wealth, expenditures, health, marriage, childbearing, child development, philanthropy, education, and numerous other topics. The PSID, though much smaller than some of the other surveys discussed here, is truly longitudinal and therefore one of the most important national sources of survey data on worker movement across occupations, firms and industries.

Researchers have identified some drawbacks for using the PSID to understand worker mobility. First, workers' occupation and industry of employment are only available annually in most cases, so if workers move more often, those transitions will not be captured, making the CPS a better choice for understanding more frequent job changes. Second, the PSID's sample, by design, excludes immigrants arriving to the U.S. after 1968 so PSID findings can be generalized only to the non-immigrant population. Finally, there is "noise" in the PSID data with regard to the actual occupation of workers in the sample, though information on occupational changes is deemed reliable (Kambourov & Manovskii, 2013).

### 3.3.3 American Community Survey

The other main survey that is representative of the country as a whole is the American Community Survey (ACS), also conducted by Census. It is a monthly survey of about 300,000 households. This contains information on the incomes and occupations of respondents, along with a lot of other data, but has no longitudinal component, as new survey participants are chosen each month. Because there is no longitudinal component, ACS is not useful for examining occupational transitions. ACS microdata are available in public use microdata sample (PUMS) files.

### 3.3.4 Occupational Employment Statistics

Through its Occupational Employment Statistics program, BLS also collects data on occupational employment and wages. BLS surveys about 200,000 establishments every six months. However, like the ACS, there is no longitudinal component; the OES survey is designed so that each establishment is surveyed at most once within a three-year period. However, because OES is such a large survey, it allows BLS to compile information on occupational wages for individual states and urban areas, and therefore is useful for benchmarking the wage outcomes of career pathways programs, as career pathways programs typically are focused on a single locality. OES microdata is not typically available, but tabulations do exist for the nation, states, and metropolitan areas, as well as individual industries.

### 3.3.5 Workforce Innovation and Opportunity Act Data

Another source of information about career progression comes from the U.S. Department of Labor's Employment and Training Administration (ETA), as part of its administration of WIOA. It is a requirement of WIOA that AJCs report information about their clients to ETA for the purpose of performance monitoring. These administrative records are referred to as Participant Individual Record Layout (PIRL) and are stored in the Workforce Integrated Performance System (WIPS). PIRL

includes information on each client's demographics, educational background, and receipt of government benefits such as Temporary Assistance for Needy Families (TANF). It also includes information on employment services and education and training received through WIOA (including the specific program) and on job placements (including whether a reported job placement is related to the training received). Each individual in WIPS is also matched to his or her Unemployment Insurance or similar earnings records for a limited follow-up period. Note that AJC clients are not similar to the population as a whole as they tend to be disadvantaged in some way (for instance, they may have recently lost a job). However, they are similar to the populations that typically are enrolled in career pathways programs, as those also commonly are targeted at disadvantaged populations.

### 3.3.6 National Longitudinal Survey of Youth

BLS also conducts the National Longitudinal Survey of Youth. This survey has been conducted in two waves, with initial surveys in 1979 and 1997, respectively. The survey is designed to follow a nationally representative sample of youth through their education and careers. For the 1997 wave, participants were born between 1980 and 1984 and were between 12 and 17 when first interviewed in 1997.<sup>25</sup> The participants are now (as of this writing in December 2017) between 32 and 37 and continue to be interviewed every other year. The survey collects information about industry and occupation of employment and about wages. Though the sample is much smaller than either the CPS or the ACS, there are many more job transitions in its database, and this represents the richest source of career trajectories data available as part of a federal government survey. However, given the smaller sample, NLSY is most useful in examining transitions between more common occupations, because less common ones are unlikely to be present in the data to an extent that allows for statistically valid conclusions about the overall frequency of transitions in the labor market. NLSY also might have limited usefulness for understanding the trajectories of young workers given the current age of the 1979 and 1997 cohorts. NLSY microdata are made available under a public use or a restricted use license from BLS; the latter license includes some information, such as geographic identifiers, that is more sensitive.

### 3.3.7 Private Company Data

Private companies are increasingly collecting large amounts of information about people, their salaries, and careers. Two of the most notable are LinkedIn (now part of Microsoft) and PayScale. LinkedIn is the leading career-oriented social network and it encourages users to put versions of their resumes on their homepages. These resumes can in principle be mined to show career trajectories, although people's occupations are not coded in a consistent manner, which makes it difficult to produce aggregate statistics. Techniques from machine learning for text classification can be employed to classify these jobs. LinkedIn does not collect compensation data, however. And LinkedIn's users are not a representative sample of the population; its users tend to be more educated.

One advantage of LinkedIn's large corpus of resumes is that it in principle represents a large longitudinal data set that would allow researchers to calculate career trajectories, although the absence of standardized titles and compensation data is an obstacle to doing this. Increasingly, though, researchers are working on ways to transform such somewhat unstructured data, such as that

---

<sup>25</sup> The survey was re-fielded in 1998 to increase sample size, so a small proportion of respondents were 18 at the time of their first interview.

found in resumes, to the types of structured data that a formal survey with a closed set of responses to each question collects.

PayScale's business model is based on providing compensation data to both employers and employees so that both parties have better information in setting and negotiating salaries. Employees provide their age, gender, race/ethnicity, job title, salary, employer, years of experience, and education (including school, degree, field, and date), as well as one prior job title. They are provided a custom salary report. PayScale then sells this crowdsourced data to employers. Because PayScale has many salaries for each job title, as well as data on prior job titles, it also has the ability to determine career trajectories.

Data from private sources such as LinkedIn and PayScale are not designed to be—nor are they—representative; but it is possible to combine the private data with government survey data that is representative. For instance, occupational transitions data from LinkedIn could be reweighted based on geography, wages, and occupational frequencies from large federal surveys such as OES, CPS, or ACS discussed above. For example, if LinkedIn's users are more educated and more urban than the workforce as a whole, then reweighting potentially could correct for this, providing that there are sufficient people in LinkedIn's database who are uncharacteristic of the database as a whole and can be reweighted accordingly.

The City University of New York's (CUNY) New York City Labor Market Information Service uses data from PayScale to construct a set of thematic career roadmaps that show such trajectories. It has produced maps for several occupations and sets of occupations: Bookkeeper and Accountant, Cook and Chef, Home Health Aide, Medical Assistant, and Tech Support Worker. Each map lays out a variety of trajectories drawn from the PayScale data.<sup>26</sup>

### 3.3.8 Structured Interviews

Structured interviews are another possibility for a data source that could be used as part of a qualitative study to understand worker's trajectories in greater depth. Interviews would be structured around common topic areas, including workers' educational and professional backgrounds; motivations, goals, and aspirations; challenges in advancing in terms of occupational steps; and the factors that workers found beneficial to their success. Interview protocols would be semi-structured to discuss the topics listed above but would be open-ended to allow study participants to go into greater depth about a particular topic.

One challenge would be to recruit study participants that are working in the industry and career level of interest. If done in conjunction with another evaluation, options could include recruiting among those who have indicated an interest in the program or recruiting from staff at employer partners. Researchers may also be able to recruit participants for such a study through relationships with large employers or by reaching out to potential job seekers through websites and other entities that advertise jobs (to the extent that doing so is allowed). While these approaches would be unlikely to yield a representative sample, they could still provide valuable information.

---

<sup>26</sup> See the CUNY Center for Urban research website for links to these maps:  
[https://www.gc.cuny.edu/lmis/information/career\\_maps](https://www.gc.cuny.edu/lmis/information/career_maps).

### **3.4 Practical Considerations and Tradeoffs**

In most career pathways programs, there is a focus on a single sector, or a few sectors. Thus, the designers and implementers of such programs are typically focused on those industries, and not the economy as a whole. Analyzing career trajectories using surveys conducted by the federal government (mainly by BLS and Census) has limitations. Such surveys do not focus on particular industries and so typically do not have sample sizes that are sufficient to produce reliable statistics about particular industries. The main exception to the rule is the Occupational Employment Statistics program discussed above, because it surveys so many establishments. However, OES does not have a longitudinal component and therefore is not helpful in understanding occupational transitions.

As discussed above, one way to handle this problem is to conduct a new survey, but that can be expensive if representativeness is the goal. Another option, again as discussed above, is to collect data from large employers in the industries of interest. Administrative data from large employers exist in proprietary databases. Though information from a single large employer can be helpful in assessing worker mobility within that employer, it is not as helpful when workers transition between employers and move from one system to another. For instance, in New York City, there are five large hospital systems: New York Presbyterian, Mount Sinai, Northwell Health, New York University, and the public Health and Hospitals Corporation. Nurses and other workers may move from one to another, as well as to smaller clinics and hospitals. Integrating data between systems poses large logistical and technical challenges. This may be less of an issue when a particular employer dominates a sector in a particular locality.

Additionally, extant data may be affected by bias stemming from the way in which the data was originally collected. This could include omitted variable bias, opt-in bias, bias in the location of the original studies that limits their generalizability, etc. The extent to which bias is a factor would need to be considered in the context of the given data set, and researchers would need to understand how the original studies controlled for such bias. Conducting new surveys could address many of these issues, but could be very expensive, as described above.

Another consideration is that because the data on occupational transitions is descriptive, it does not allow answers to causal questions, such as what factors influence transitions into (or out of) a particular occupation. As discussed above, individual fixed effects models have the potential to address the impact of particular programmatic elements, or other interventions, on these transitions; but these models require longitudinal data on each individual involved, as well as administrative data on the intervention. This research approach does have the potential to make fairly rigorous estimates of impacts, even outside the context of an experimental design.

Maps of career trajectories are most useful in building understanding of the current structure of the labor market, which can inform career pathways program design. Qualitative studies based on in-depth interviews with workers could go a step further in informing career pathways program design or in illuminating trajectories in fields that are poorly understood, although they are unlikely to be representative and their limited generalizability should be recognized. They also do not help in estimating program impacts, unless the data on the transitions are coupled with a research design such as an experimental evaluation or an individual fixed effects model. In one of these contexts, however, differences in occupational flows between treated and untreated individuals can themselves serve as impacts of interest.

## 4. Role of the Public Workforce System

This chapter focuses specifically on the role of the public workforce *system* in career pathways and the priority research question on this topic that emerged from the project’s knowledge development phase. We examine the workforce system through the lens of AJCs and their relationship with career pathways initiatives. This chapter also introduces two main lines of inquiry about the workforce system through an implementation research framework. After introducing and discussing the relevant research question in Section 4.1, Section 4.2 discusses three approaches for addressing it: reviewing existing studies; conducting descriptive, implementation or other qualitative research; and pilot studies. Section 4.3 describes relevant data sources, and Section 4.4 discusses practical considerations and tradeoffs.

### 4.1 Introduction to Workforce System Research Question

The Career Pathways Design Study has one primary research question considering the public workforce development system and career pathways approaches:

**What roles in career pathways initiatives is the workforce system playing as compared with other entities that more commonly lead these interventions, such as community and technical colleges?**

In the context of career pathways approaches, the *workforce system* is the collection of publicly funded institutions and structures that offer training, career advisement, and employment benefits for job-seeking individuals. These institutions primarily involve AJCs, which are located within networks of local and regional Workforce Development Areas (WDAs) governed by local and state Workforce Development Boards.

The research question in this chapter addresses the extent and nature of the workforce system’s role in career pathways initiatives at the system and program levels. Career pathways initiatives commonly have been led by community and technical colleges (34 out of 52 studies from the *Research and Evaluation Synthesis* reported one of these institutions as the main site), though the public workforce system is often a partner (16 studies). As an extension of the primary research question, we also examine research approaches that could address the following related questions:

- To what extent is the workforce system’s role different at the program and system levels of career pathways initiatives? Are these roles shaped by the structure of past federal grant initiatives and workforce policies? How might they change with the implementation of WIOA and new grant opportunities?

In examining these related questions, it will be important to consider the role of the workforce system in career pathways through its partnerships and coordination with career pathways training providers as distinct from the community and technical college system’s role in career pathways initiatives. For example, one role played by AJCs is to fund career pathways training at community and technical colleges through Individual Training Accounts. Other roles for the workforce system may include assessing participants, referring them to a range of benefits and supportive services other than training, and helping place participants in jobs.

## 4.2 Research Approaches

This subsection discusses three approaches for answering questions about the public workforce system’s role in career pathways initiatives: reviewing existing studies, qualitative research, and pilot studies.

### 4.1.1 Review Existing Studies

One research option is to conduct a review of existing studies on career pathways initiatives with a specific focus on the AJCs. Although many of the impact evaluations listed in Exhibit A.1 involve AJCs to some degree, such as the Green Jobs and Health Care Impact Evaluation, Pathways for Advancing Careers and Education, and Accelerating Opportunity, there are other studies that evaluate them qualitatively. Much of the previous work on AJCs (and previously, One-Stop Centers) focused more on features of the system (e.g., Institutional Analysis of American Job Centers<sup>27</sup>), and less specifically on their role in career pathways initiatives at the system and program levels. If a thorough review of the literature confirms that this is indeed the case, then more research on that role could be worthwhile.

### 4.1.2 Descriptive, Implementation, or Other Qualitative Research

A second option is for DOL to make this research on the workforce system role a part of implementation studies on career pathways strategies as described in Section 2, and start with interviews with selected workforce officials at the state and local levels, including AJC managers and staff. This initial implementation research could focus on the following:

- To what extent do the state and local workforce agencies view it as their role to help promote or support career pathways programs? What types of activities do they provide (e.g., case management, career guidance, education and training referrals, financial assistance)? To what extent is implementation of career pathways strategies by the workforce system part of the goal of DOL grant programs and/or initiatives at the AJC, local WDA, and/or state levels?

These questions would identify current efforts and priorities and provide insight on the potential areas for growth, as well as challenges the system faces in involvement in career pathways initiatives.

Framing these implementation research questions could take two lines of inquiry. In the first, these implementation questions could be asked within the context of broader impact studies or outcome studies about career pathways approaches. These questions would help supplement the findings of quantitative analyses on the possible effectiveness of career pathways strategies, with information on how and why the workforce system may or may not have supported career pathways programs. This option could include interviews and focus groups with local employers, to determine their relationships with local AJCs, their involvement in career pathways programs, and the extent to which the AJCs facilitate the involvement of employers in career pathways programs.

In the second line of inquiry, DOL could examine the extent to which career pathways are being implemented by the workforce system in the way envisioned by WIOA. This is a broader framework of implementation, with a focus on the implementation of WIOA at the local level and its experiences

<sup>27</sup> See description of the project here: <https://www.mathematica-mpr.com/our-publications-and-findings/projects/institutional-analysis-of-american-job-centers>.



in establishing career pathways programs. This examination would involve considering explicitly the extent to which the workforce system is playing the role that WIOA directed. It also could help differentiate the role that AJCs play in career pathways programs that may occur as part of a DOL or other grant program<sup>28</sup> versus the role that AJCs play as part of the regular course of implementing WIOA. Such an analysis would allow an examination of this question:

- To what extent do AJCs see the WIOA policy environment (federal performance measures, funding levels, and state/local policy signals) as being supportive of a career pathways approach?

Within this framework, the study could examine the extent to which the transition to WIOA has focused more attention on career pathways strategies, or elements of these strategies such as sector training partnerships or integrated basic education and training.

### 4.1.3 Pilot Studies

A third option is to conduct pilots of program interventions within the workforce system that specifically focus on promoting movement along a career pathway. The goal of these pilots would be to test ways to increase AJC involvement in career pathways efforts. Pilot approaches to test could include AJC staff actively promoting progression along a pathway through regular meetings and reminders that focus on career advancement as a goal, or they could include having AJC staff refer clients to state or local career pathways initiatives and then place them in jobs once training ends or help them transition to the next step on the career pathway. Pilots also could include enhanced access to skills training and supports with connections to jobs along a career pathway, including career navigation services to guide workers' training choices. Much of this research would include qualitative data collection and analyses from interviews with state and local workforce staff, as well as from employers and training providers. It also could include quantitative data collection and analysis of clients' baseline characteristics, service receipt, and employment outcomes. Depending on the pilot intervention tested, more rigorous experimental studies could be designed (see Section 2 for a discussion of this approach).

## 4.3 Data Sources

This subsection discusses two data sources for workforce system research—primary qualitative data and administrative data—and addresses potential data concerns.

These research approaches on the workforce system and career pathways should start with interviews of AJC staff and clients to identify and refine questions for the two lines of qualitative inquiry and identify the current state of AJCs and their relationship with career pathways initiatives. An online survey of the AJCs is also possible to gather more systematic data on the connection to and challenges and successes with career pathways programs. DOL could also carry out this type of qualitative research with local employers, in order to understand the AJC-employer relationship in the context of career pathways initiatives.

---

<sup>28</sup> For example, DOL has funded career pathways approaches and evaluations through such initiatives as the Trade Adjustment Assistance Community College and Career Training (TAACCCT) and Workforce Innovation Fund (WIF) grants. The U.S. Department of Health and Human Services' Administration on Children and Families has supported career pathways strategies and evaluations through the Health Profession Opportunity Grants (HPOG) and the Pathways for Advancing Careers and Education (PACE) evaluations.

There are several administrative data sources that can provide descriptive information on this relationship between the workforce system and career pathways initiatives, as well as provide information on potential outcomes.

First, educational records, ideally administrative data from the community colleges (or other training providers) themselves, would be best for keeping track of participants' progression through career pathways training programs. These data would include full course transcripts to track their enrollment and completion of specific course requirements, as well as information on the credentials earned (e.g., certificates, associate degrees). Demographic and assessment information, such as educational histories and test scores, also are potentially available from state community college data warehouses. This information could be supplemented with information from the National Student Clearinghouse, in case students transfer to four-year programs or attend another school from which data are not being collected directly.

In addition, data on demographic characteristics and service receipt are available through the Workforce Investment Act Standardized Record Data (WIASARD), and now the PIRL. The PIRL data offer information on individuals who receive services from AJCs, such as WIOA-funded training and other resources; they also provide information on background characteristics, as well as educational and wage histories. Wage outcomes are tracked for up to a year after AJC exit. To supplement this wage information, UI and NDNH data on these AJC clients would be important to track their long-term wage and employment outcomes. Additional wage and employment data, as well as federal benefits received, can be obtained in Social Security Administration or IRS data.

With any of these data sources, there are two main concerns: the ability to link individual-level information across the various data sets; and privacy and consent issues, as the data were collected through different systems. To be able to fully track career pathways program progress, from AJC entry to course enrollment and wage outcomes, those linkages have to occur. Linkages must be made in ways that privacy and consent procedures are considered. At the moment a lack of integrated data sources hinders this type of research as it is generally not possible, for example, for researchers to obtain information on which specific community or technical college programs or courses are being supported financially by WIOA at the individual level. Such information would be very valuable in understanding career pathway program progress and the role of the public workforce system. DOL's Workforce Data Quality Initiative grants are aimed at addressing this issue by helping states to match education, training, and employment data to evaluate workforce development outcomes and develop consumer information tools.

#### 4.4 Summary

There are various issues to consider when focusing on the workforce system in its current form, and proposing new research ideas on the system and its relationship to career pathways initiatives. The key concerns are determining whether and how the workforce system is engaged in career pathways work, and how it aligns with WIOA expectations. These considerations will need simply to be part of the analysis that addresses the research questions of interest.

The options we presented here for answering these research questions include implementation research that will help DOL examine the current state of such engagement, successes and challenges, and strategies for increasing involvement in the future.

## 5. Cost Analyses

When a policymaker or program director reviews the details of a particular program or component for potential adoption, cost is a crucial dimension in the decision-making process. The relationship between units of outcome and costs (e.g., how many people received training services per dollar spent) or monetary benefits and costs (e.g., what was the change in earnings per training dollar spent) provides a yardstick by which specific programs or components of a program can be compared. For example, which is more desirable: a costly intervention that produces very large average impacts or a more affordable intervention that produces more moderate average impacts? Thus, cost alone is a less informative dimension than how costs relate to the benefits of a program.

The measurement of all pertinent costs and benefits can be complex, and this chapter focuses on specific approaches for building policy-relevant and informative research in this domain.

### 5.1 Introduction to Cost Research Questions

Relatively little research has considered the costs associated with operating career pathways programs. As a result, the following is of priority interest for future research:

What are the costs associated with a given career pathways program to participants, funders, and employers, including opportunity costs? What do those costs “buy”?  
And to what extent are the benefits worth the cost?

A variety of costs are associated with career pathways programs, and these costs are borne by numerous actors, including participants, taxpayers, and society, including employers. The cost of the *inputs* to establish and operate the program (e.g., staffing, materials, curricula development, creation and maintenance of partnerships, tuition assistance, support services, facilities, administration, overhead) is covered by the funders (taxpayers, if using public dollars; society, if using private dollars). Program participants also bear a cost by attending the program (even those who may receive tuition assistance) in the form of unreimbursed expenses, plus the *opportunity cost* of attending the program, reflecting what participants have given up to attend, such as wages. Employers may bear costs related to helping design programs and curricula so that they meet industry needs, offering work-based learning opportunities and paying for training for incumbent workers.

One evaluation option may be to understand and characterize only these costs, rather than relating them to benefits. For example, how does the cost per outcome (or cost per person) compare with other similar programs? Other studies may be interested in comparing these costs with potential benefits; in this case, a conclusive measure of benefits (*impact estimate*), is necessary.

Benefits also are distributed among a range of actors: Participants earn increased wages and possibly also gain job benefits, such as health insurance and paid leave; public benefit programs such as TANF require less funding from taxpayers; employers hire more productive workers or see improved outcomes for clients; and society realizes the benefit of a more educated population or an improvement in worker productivity. Specific benefits to employers beyond higher productivity may include reduced job vacancy rates and lower recruitment and hiring costs. However, not all benefits are measured in dollars and cents; thus, to compare them with costs, researchers need to determine the monetary value of potential benefits. Furthermore, benefits may not be realized immediately (e.g.,

earnings gains accrue over a worker’s lifetime), so this analysis must include an approach to comparing all of the costs with all of the benefits (both current and future).

Finally, cost analysis should result in an expression that allows the result to be compared against other programs. If the focus of the analysis is only on cost, this likely will be some measure of *average cost* (e.g., average cost per person). If the analysis compares costs with benefits, this measure will indicate whether the costs outweigh the benefits, or vice versa.

## 5.2 Research Approaches

Analyses of cost generally fall into one of three categories: *cost allocation analysis* establishes how much resource is required for a given measure of intensity or magnitude of services; *cost-effectiveness analysis* establishes how much resource is required for a given change in outcome; and *cost-benefit analysis* compares dollar-valued inputs with dollar-valued outcomes. The type of cost analysis dictates the comparison that can be made with other programs. Cost-effectiveness analysis can demonstrate which programs produce desired outcomes for the lowest cost, whereas cost-benefit analysis results in a clear policy recommendation when comparing across programs—the lowest cost per benefit is the program that should be adopted.

The most appropriate cost analysis for a research study will be driven, in part, by the research design. Designs that produce impact estimates, such as experimental evaluations, are good candidates for a cost-benefit analysis because the outcomes and costs of the control group provide the necessary context for the outcomes and costs of the treatment group. That is, the benefits are calculated as the value of outcomes for treatment group members less the value of the corresponding “business-as-usual” outcomes for the control group members. It is important that this calculation should not double-count the opportunity cost of foregone earnings; this feature is implicit in the earnings impacts. Similarly, costs are calculated as the value of all resources used to provide the program to the treatment group less the value of any resources to provide the “business as usual” condition to the control group.

A cost-benefit analysis is not appropriate for designs that do not produce impact estimates because the monetary value of the program benefits cannot be determined and the counterfactual (“business-as-usual”) condition remains unknown. In this case, researchers may choose to collect information on costs incurred by programs and calculate costs per participant or costs per participant outcome; that is, a cost-effectiveness analysis. Researchers also can relate these costs to the sources that were used to fund the programs (e.g., public versus private dollars) (Holl et al., 2009). Cost allocation and cost-effectiveness analyses should make an effort to provide context for their findings by comparing them with cost analyses for similar programs, either through their own analysis of those other programs or from the literature.

As detailed below, the key inputs in a DOL career pathways study would be costs associated with each career pathways program itself (e.g., staffing, materials, tuition assistance, support services, administration, overhead) and the added costs of subsequent education and training (e.g., foregone wages, tuition, public subsidy of higher education) that result from participation in the program. The hypothesized impact of each career pathways program is increased education and training for its program participants, which is expected to result in improvements in employment and income. The most appropriate approach for measurement is likely to vary for inputs and outputs depending on data

availability, the form of findings, and the form of estimates taken from the relevant literature to make necessary model and parameter assumptions.

Note that unlike previous research approaches discussions elsewhere in this report, the rest of this subsection describes what amount to steps in a cost analysis process (rather than a selection of stand-alone research approach options for DOL's consideration).

### **5.2.1 Identifying the Costs Associated with Career Pathways Programs from Various Perspectives**

The costs of implementing, operating, and participating in a program are borne by numerous actors with various perspectives. The analysis of costs should separately aggregate costs for various specific groups, such as the following:

- **Program participants.** Costs to treatment group members include reduced earnings and available time (e.g., for family commitments or leisure) during training. The largest cost of program participation from the treatment group member perspective is the opportunity cost of foregone employment (and thus earnings) during that participation. Analysis of opportunity costs should calculate the extent of foregone employment at each site by direct contrast with the control group at the site. Specifically, impact models can be used to estimate differential earnings for the treatment group relative to the control group during the training period after random assignment. This difference (anticipated to be negative) represents the foregone employment cost of the intervention.
- **Taxpayers.** Many programs are funded by federal and state/local governments; in these cases, taxpayers incur tangible costs such as facilities, staffing, and materials. The most straightforward aspect of an analysis of cost is collecting data on tangible costs incurred by program funders. A standard approach for gathering these data is to collect them from programs directly, either via survey or by collecting and reviewing administrative documents that contain cost information. An additional dimension of taxpayer costs relates to the marginal social cost of public funds. Each dollar of government spending, (e.g., to fund a career pathways program or to provide a social support like the Supplemental Nutrition Assistance Program (SNAP)) costs society more than a dollar due to the cost of operating and distortions induced by the tax system.
- **Society.** When programs are not funded by the government, society bears the cost of program funding. The approach for collecting these cost data is the same as for government-funded programs. Businesses that employ program participants prior to their entering the program suffer from loss of workers or reduced productivity if participants replace work with training. For our purposes here, employer costs are included as part of society's, but one could develop a study that broke their costs and benefits out separately. Identification of costs from the perspective of employers is very challenging; researchers can use targeted surveys of employers in specific industries or extant data sources to understand changes in employee hours, turnover, and productivity.

To organize and summarize the costs across different perspectives, an exhibit such as Exhibit 5 can be populated with program-specific information. For demonstration, the exhibit includes categories and components of costs that are likely to be included in career pathways programs.

**Exhibit 5. Expected Costs of a Career Pathways Intervention, by Perspective**

Category	Component	Expected costs of a career pathways intervention, by perspective			
		Program Participants/ Treatment Group (relative to control)	Taxpayers	Employers	Society
Unique program elements	Pre-college or program-specific non-mainstream coursework	Opportunity cost of time (foregone earnings)	Funding	Loss of workers; reduced productivity	Funding
	Direct assistance (cash and education materials) to treatment group members	No change	Funding	No change	Funding
	Program administration	No change	Funding	No change	Funding
Mainstream education or training	Education or training not unique to the program	Opportunity cost of time (foregone earnings) and cost of tuition	Funding for higher education	Funding for higher education (if offered)	Funding for higher education

**5.2.2 Assigning a Monetary Value to Benefits for Participants**

As with costs, whether or not the outcomes of career pathways programs justify their costs depends very much on one’s perspective. Both this subsection and Subsection 5.2.3 describe approaches to measuring benefits across differing perspectives.

In general, program costs are incurred within the first few years of program participants’ enrollment, whereas potential benefits accrue through increased earnings for program participants over their lifetimes. In the career pathways framework, program participation may result in educational attainment for multiple years after random assignment, as students complete a program, become employed, return to complete a higher-level credential, etc. It is possible, depending on the local theory of change, that gains in earnings may not materialize for some time after program participation. For this reason, analysis of benefits must consider the value of future benefits, in some instances projected beyond the observed data, and always discounted to current dollars for comparability with costs and benefits that are realized today.

Projections and net present value calculations require making assumptions about the future earnings profiles of program participants based on outcomes observed during the timeframe of the study. To make the necessary projections, the analysis should use published literature on the returns to education for low-income workers, as well as long-term follow-up of evaluations that occurred sometime in the past. Previous rigorous studies, such as Bell and Orr (1994), offer approaches for developing lifetime earnings profiles based on observed earnings shortly after random assignment. More recent work extends this framework to consider other domains in addition to lifetime earnings, such as health, lifetime criminality, and educational attainment (Garcia et al., 2017). Other recent literature, particularly studies that focus on similar target populations, may have additional insight into projecting lifetime earnings profiles. Analysis of earnings can also be further decomposed into a wage component and an hours component (though wages and hours typically can be measured only through survey questions); understanding whether observed changes are due to shifts in hours or wages can play an important role when projecting a lifetime earnings profile.

Last, there are individual benefits other than employment-related outcomes that should be considered for inclusion. For instance, education is often thought of as providing nonmarket personal benefits of greater life satisfaction; employment offers fringe benefits, such as health insurance and paid sick leave; and children of treatment group members may benefit from improved education and economic

outcomes. The analysis of benefits should clearly define the benefits that will be monetized and included in the analysis. As with projecting earnings trajectories, prior literature can describe strategies for monetizing these specific benefits. Certain benefits may be too difficult to monetize and include in the analysis as to society; these benefits can be labeled as secondary and described separately.

### 5.2.3 Measuring Benefits to Taxpayers and Society

The impacts of career pathways programs are not limited to participants. Favorable impacts for participants also are likely to extend to taxpayers and society.

Taxpayers can realize benefits from a career pathways program in the form of changes in taxes, public benefits, and crime/incarceration. Some of the changes to public benefits and taxes represent changes in transfers that work in opposite directions. For instance, when training results in increased skills and employment in higher-paying jobs, those participants may receive less public assistance and pay more in taxes while the government pays less in public assistance and receives more in taxes. For some populations, these transfers may entirely cancel each other out (except for a small savings in administrative costs). However, even though changes in benefits and taxes may be close to neutral from the societal perspective, such changes can make a big difference from a participant perspective and the taxpayer perspective, and thus are important to a cost-benefit analysis. In addition, the decreased administrative costs associated with such transfers do represent a net positive gain from a societal perspective.

A number of resources exist for estimating the effects to taxpayers of changing conditions, such as increases in income. The net effect of increased income on benefits received and taxes paid is referred to as the *effective marginal tax rate*. One strategy for incorporating the effective marginal tax rate into analysis is to use estimates compiled by the Congressional Budget Office (CBO). The CBO conducts in-depth analysis on the effective marginal tax rate for low- and moderate-income workers. That analysis can be applied to the estimated lifetime earnings profiles in order to estimate the total reduction in public benefits plus the increase in taxes paid that accompany earnings gains (Congressional Budget Office, 2015).<sup>29</sup> Though the CBO analysis includes some state tax and transfer payments, additional resources can be used to be sure all state and local tax information is captured. The Center for State Tax Policy at the Tax Foundation publishes information on state and local tax rates that can be used to supplement the CBO analysis as necessary in estimating changes in state and local tax revenue that will result from estimated changes in lifetime earnings.<sup>30</sup> Also, researchers can try to model any potential multiplier effects on the local economy if they anticipate the injection of earnings to result in additional tax revenue that is not localized to study participants.

Reductions in benefits receipt due to increases in earnings will result in decreased administrative costs associated with the benefits programs. These reductions are a benefit to taxpayers. A common approach to monetizing this benefit is to assume an average level of administrative costs across all public benefits programs and to apply this multiplier to the estimated change in public benefits receipt (Karoly, 2008). Studies typically assume around 10 percent, which is in line with most estimates of

<sup>29</sup> Note that this requires an assumption that current benefits and tax policy remain in place. However, this analysis is an efficient approach in that it does not repeat the exercise already completed by the CBO of estimating changes in benefits for TANF, Medicaid, UI, SNAP, housing assistance, and other benefits individually.

<sup>30</sup> Center for State Tax Policy [website]. Accessed October 28, 2017. <http://taxfoundation.org/tax-topics/state-taxes>.

administrative costs found when examining budget details for the relevant federal and state agencies (Karoly, 2008; Isaacs, 2008).

Reductions in crime and incarceration are benefits to taxpayers, who fund police forces and many prisons/jails, as well as to society; these benefits can be measured directly. The benefits to taxpayers associated with such reductions can be calculated using estimated costs of fighting crime, prosecution, and incarceration (e.g., Cohen, 2004). The benefits to society are more difficult to measure, but include the economic growth and social stability associated with lower overall crime rates. As with benefits to participants, these benefits can be described separately if they are too difficult to monetize.

Last, the subset of society that are employers can realize benefits through two primary domains: quality of labor and productivity. Suppose employers are struggling to fill currently vacant healthcare positions because they require applicants to have certain credentials, and employers lack the foresight to understand that an investment in career pathways programs will produce credentialed, skilled workers in the future. In such a case, if career pathways programs are increasing the average skill level of program participants, employers may find it easier to meet their demand for skilled workers. This creates the potential for a stronger match between employer and employee, leading to a reduction in turnover. The employer may also realize benefits of the career pathways programs through improved productivity in the production of goods or provision of services—increases in productivity will be realized in the employer cost savings.

#### **5.2.4 Relating Costs to Benefits**

Once costs (and, if included, benefits) have been identified and measured, the researcher must choose a methodology to express the results of the analysis. As described in the preceding subsection, impacts of increased employment and income for program participants will change the public benefits received and taxes paid by them over their lifetimes. Because benefits and costs accrue over a long period of time and at varying intervals, monetized costs and benefits should be discounted to the year treatment group members enrolled in the program before the costs and benefits are compared, such that the comparison determines the net *present value* of each program. This present value approach allows an “apples to apples” comparison of costs and benefits that occur at different times.<sup>31</sup>

For cost-effectiveness studies, the most appropriate result to report is the average cost per participant or average cost per completed training. For cost-benefit analyses, several possibilities exist for comparing monetized inputs and outcomes (Dolfin & Schochet, 2012). One popular option is subtracting the present value of costs directly from the present value of benefits (the resulting value is also known as the *net present value*, or NPV). Another option is calculating the ratio of the present value of benefits to the present value of costs. Generally, ratio measures relating benefits to costs are easy to understand and interpret. However, these measures are sensitive to the classification of costs versus benefits. For instance, if career pathways programs reduce the costs of pursuing additional training, ratio-based measures will be sensitive to whether this value is added to program benefits or

---

<sup>31</sup> Calculating the present value requires an adjustment of all monetized costs and benefits to the same year by applying a discount rate. For example, an researcher might follow the Office of Management and Budget (2003) guidance in Circular A-4 and use a discount rate of 3 percent, which is recommended for a program that “primarily and directly affects private consumption” where a “social rate of time preference” is the appropriate adjustment. Once a discount rate has been selected, monetized benefits based on literature estimates also should be adjusted accordingly.



subtracted from program costs (Dolfín & Schochet, 2012). Therefore, the remainder of this section assumes researchers relate costs to benefits using the NPV.

In addition to the overall NPV, researchers also can recalculate NPV according to the perspectives identified earlier. Specifically, from a *treatment group member perspective*, the NPV indicates whether enrolling in the program makes sense as a personal investment. From a *taxpayer perspective*, the NPV indicates the net taxpayer resources (adjusting expenditures today to account for lower expenditures and higher tax revenue in the future) needed to generate benefits experienced by taxpayers, such as reductions in crime and incarceration. Finally, from a *societal perspective*, the NPV indicates the net social resources needed to generate benefits experienced by segments of society, such as employers.

In addition to these cost analyses by perspective, research can attempt to disentangle the costs and benefits of various program components. Such an analysis could be particularly fruitful in identifying which components generate the largest returns in excess of their cost. Detailing the costs of different program components might be straightforward if cost information for the program component of interest is available at the proper level of granularity. However, this calculation could be complicated if those data are not available or there are costs that are common to various program components (Barnow & Trutko, 2010). In such a case estimating the cost of a program component may require researchers to make assumptions about how total program costs are distributed across program components. In addition, the ability to precisely detect the benefits due to those specific components can be very challenging. Such an analysis might be feasible in a career pathways program that randomly assigns specific program enhancements such that the benefits of the enhancement can be experimentally estimated.

Finally, research can expand the analysis to understand how costs and benefits differ by observable subgroups of study participants. For example, differences in the costs and benefits of the program for participants who enter with certain levels of educational attainment or those who enter from employment versus unemployment might be of particular programmatic interest.

### 5.3 Data Sources

A variety of data sources are likely to be available for a cost analysis of any career pathways program. This section summarizes different types of data sources:

- ***Profiles of programs*** give a high-level overview of the program, plus detail program goals, target population, and structure. Profiles also describe each program’s career pathways components. This information can be used to identify program inputs. It also can be used to prepare survey or cost collection interview questions.
- ***Qualitative data from an implementation study*** such as notes from site visits, documents related to the program, and observation of program activities provide details on program components and structures that might be useful to prepare survey or cost collection interview questions. This information also can be used to estimate costs if direct data collection is not feasible.
- ***Cost data collection interviews*** are interviews with program staff specifically focused on gathering information on operations and costs. These interviews allow researchers to develop a comprehensive list of program inputs and associated costs.

- ***Follow-up surveys*** are designed to collect information on outcomes such as service receipt, educational attainment, income, debt, and participation in public benefits programs. As described earlier, such outcomes can be important for measuring benefits and costs.
- ***National administrative data*** are a second source of reliable information on outcomes. Specific data sets, such as those described in Exhibit A.2, might offer access to comprehensive records of earnings, employment, and educational pursuits, among other possible outcomes, for everyone involved in the study, whether in the treatment group or not.
- ***Program records*** might only have information for treatment group members (versus follow-up surveys or administrative data, which could have data on both treatment and controls), but program records could be useful to refine estimates of costs to individuals. For example, a more precise measure of the cost to degree/certificate attainment in terms of specific program components might make use of program records.
- ***Surveys of employers*** can be used to identify benefits to employers of the program, particularly if the intervention is targeting a particular industry and/or has partnerships with employers such that identification of employers is feasible.

A logical starting point for identifying the most appropriate data sources for a career pathways program evaluation is the set of studies summarized in Exhibit A.1 as well as the data sources summarized in Exhibit A.2. Several of the studies in Exhibit A.1 have included cost analyses, summarized in Exhibit 6, and DOL can use those studies to learn about techniques and comparable findings.

**Exhibit 6. Cost Analyses in Completed and Ongoing Studies**

Type of Cost Analysis	Study
Cost analysis	<ul style="list-style-type: none"> <li>• WorkAdvance</li> <li>• On-Ramps to Career Pathways</li> </ul>
Cost-benefit	<ul style="list-style-type: none"> <li>• Health Profession Opportunity Grants (HPOG) 2.0 National and Tribal Evaluation</li> <li>• JOBSTART Demonstration</li> <li>• Minority Female Single Parent Demonstration at the Center for Employment Training</li> <li>• Pathways for Advancing Careers and Education (PACE)</li> <li>• SNAP E&amp;T Pilots</li> <li>• Accelerating Opportunity</li> <li>• I-BEST (Community College Research Center I-BEST Evaluation)</li> </ul>
Cost-effectiveness	<ul style="list-style-type: none"> <li>• Accelerated Training for Illinois Manufacturing</li> <li>• Summer Career Pathways Project</li> <li>• YouthBuild Evaluation</li> </ul>

**5.4 Practical Considerations and Tradeoffs**

Cost analysis should use observed study participant and program data wherever possible to measure impacts and outcomes, but complete data on all benefits and costs are not always available. If data are not available, imputation techniques and assumptions regarding values for costs or benefits can be used to complete the analysis. This includes taking benchmarks from external research and developing projections about future earnings based on observed data.

Additionally, the results of cost-benefit analysis can hinge on assumptions and projections such as those described in the prior subsections. To assess how important these assumptions are to the conclusions, researchers should follow the standard practice of conducting a sensitivity analysis for each assumption. Specifically, NPV should be recalculated using alternative values for the various assumptions made.

## **6. Summary and Framework for Considering Research Options**

This report provides a set of options for DOL to consider for future research on career pathways initiatives. The research questions discussed in this report stem from a knowledge development effort (completed in February 2017) that produced reports on (1) research and evaluation relevant to career pathways approaches, (2) the implementation of existing career pathways initiatives, and (3) the potential for career pathways approaches in early care and education (ECE). Through that work we examined existing research on career pathways, identified gaps in knowledge, and solicited input on high priority questions for the career pathways field. We then narrowed down the questions in collaboration with DOL to a list that reflect those most relevant to the workforce development field and that also are less likely to be fully addressed by current ongoing research. It is important to note that a number of ongoing studies have reported or will soon be reporting early impact and implementation results subsequent to that knowledge development effort; DOL may wish to revisit their career pathways research agenda as results are released so that future research can be informed by the very latest knowledge.

The report has considered several types of impact research questions. One of these involved the long-term follow-up of career pathways program study participants. Another involved the study of individual program components in the context of an experiment. A third involved examining impacts on subgroups of interest. We also examined research questions that involve the description of workers' career trajectories in the labor market in general and in particular industries, outside the context of career pathways programs, as well as a research question that would explore the role of the workforce system in these programs. Finally, we examined cost-related questions. For each question, we identified one or more possibilities for answering it, along with identifying potential data sources. Across these, we have considered qualitative and descriptive approaches as well as quantitative approaches that can help advance the state of knowledge; indeed, bringing multiple perspectives together will enrich what we can learn about the questions at hand.

This concluding section synthesizes the wide range of information and options contained within this report to assist DOL in selecting the most appropriate design option(s) for future investments. Specifically, it focuses on the nature of each research question (that is, *what* to ask: causal or descriptive), considerations in selecting different options (that is, *how* to ask: what approaches to take); and factors in choosing research locations (that is, *where* to ask). The remainder of this chapter elaborates on these what, how, and where aspects of future research.

### **6.1 What to Ask: The Causal or Descriptive Nature of a Research Question**

The first step in implementing one of the design options recommended in this report is to clearly characterize the nature of the motivating research question as descriptive or causal. Earlier sections of this report demonstrated that this is directly related to the research design. For example, while an experimental design may be necessary to address a causal question, it is not necessary for a descriptive one. The selection of what research questions to pursue will most likely be informed by the variety of ongoing studies at DOL and elsewhere.

## **6.2 How to Ask: Approaches to Research Questions and Framework for Considering Options**

This section explores considerations for choosing among methods for approaching the four groups of research questions for future studies of career pathways initiatives. The nature of various design option tradeoffs differs depending on the nature of the motivating research question, and the specific research question itself. We describe these considerations separately for questions focused on measuring causal impact, implementation-related questions, and questions focused on addressing career trajectories and cost research.

### **6.2.1 Approaches to Addressing Causal Impact Questions**

Below we discuss the rigor of various approaches to answering questions of causality. In the context of research to identify the causal impacts of an intervention, rigor has two main dimensions: internal and external validity. As noted in Chapter 2, internal validity refers to an evaluation design's ability to rule out alternative explanations for an observed change in outcomes. External validity refers to the generalizability of the evaluation's results: the results have relevance to other places, populations, and times.

Experimental designs are the most rigorous in terms of internal validity, their ability to answer causal questions: they allow for the discovery of causal relationships with a high level of certainty. Quasi-experimental designs are less rigorous for answering causal questions, and comparative designs that do not attempt to mimic experimental conditions are the least rigorous. As for the dimension of rigor that considers external validity, quasi-experiments are commonly thought to have the potential to be better than experiments. That said, current research on the external validity of experiments suggests that this need not be the case.<sup>32</sup> Indeed, future experiments should implement the recommendations from that research in order to enhance the potential generalizability of their results. For this reason, we suggest that, on external validity grounds, there is not a large difference among the approaches considered here. As a result, the rigor assessment (see Exhibit 7) focuses on internal validity and the approach's ability to support causal claims.

It is beyond the scope of this report to examine specific evaluation costs; and any specific research would involve a wide variety of factors that would contribute to its cost. Instead, we provide general estimates of whether costs of an approach are anticipated to be low, moderate, or high. A rough gauge for these might be that a "low" cost effort can reasonably be undertaken for less than \$1 million, and certainly in some cases much less than that. A "moderate" cost effort would be in the range of \$1-2 million, and a "high" cost effort would cost more than that, possibly being a multi-million dollar investment.

We recognize that important efforts are underway to encourage low-cost experimentation (Laura and John Arnold Foundation, 2015). These efforts achieve low cost by involving existing initiatives where key outcome measures come from existing or administrative data. We urge DOL to consider how those recommendations can be integrated with future research planning. That said, not all questions can be answered within those constraints, and it will cost more to undertake new

---

<sup>32</sup> Peck (2017) explains the issues and links to the major sources in this literature.

demonstrations, try new programmatic advancements, and collect data via surveys when administrative data do not capture important outcomes of interest.

As shown in Exhibit 7, the most rigorous approach for addressing questions of causality would be to conduct new experimental impact studies. Depending on the specifics of the evaluation (e.g., whether outcome data can come from existing, administrative sources or would require a survey), new experiments are probably also the most costly of the options for examining causal impacts. We consider this option to be of high value because DOL would be able to determine all aspects of the study, shaping the research to meet its very specific information needs.

**Exhibit 7. Framework for Considering Tradeoffs across Research Approach Options for Causal Questions**

Approach	Rigor	Cost
Conduct new experiments	High	Moderate-High
Re-analyze existing experiments	High	Low-Moderate
Use quasi-experimental designs, analyses	Moderate	Low-Moderate
Meta-analyses	Moderate-High	Low

Generally, approaches to impact estimation that are less costly and easier to carry out also are those of lower rigor, although this is not an absolute rule. Reanalyzing existing experimental data may be a relatively low-cost option, but DOL would be limited in its questions to those that would be answerable given the existing data. Finally, a meta-analysis across existing studies—especially if those selected for analysis used rigorous designs—can provide high-value and rigorous results, at a relatively lower cost.

Consider the case of analyzing the impacts of a career pathways initiative on a given participant subgroup of interest. Analyzing the impacts of a program for a subgroup that was not previously analyzed in an existing study is relatively straightforward and can be done (once data are available) relatively quickly at a low cost. Mounting an entirely new study will take more time, effort, and resources. That said, the new study might answer a greater number and diversity of research questions beyond the specific subgroup of interest, and so the greater investment might be worthwhile. In addition to answering causal questions, answering descriptive questions is also important; and this project has identified them as among the high priority areas of study.

**6.2.2 Approaches to Conducting Implementation Research**

Like other research designs, the approach to an implementation study is driven by the research questions and can be used to meet different goals. Implementation research can be used to support impact studies, providing critical qualitative and quantitative information for generating theories and research questions, interpreting impacts, and replicating results. As such, implementation research is an essential partner to these studies in that it provides important details about the content and operations of the program or initiative being evaluated and the extent to which the program operated as planned. Qualitative implementation research can also supplement the understanding of other quantitative analyses, such as gaining the worker perspective on factors that affected program participation or career trajectories, to shed light on potential causes of observed patterns. In these cases, the costs of implementation research are part of a larger, often mid- to high-cost study. Relatedly, implementation evaluation can often serve as a gateway for decisions about impact studies

(e.g., Epstein & Klerman, 2016). Ideally, only once a program is fully and faithfully implemented is it ready to have its impact evaluated.

Implementation studies can also stand on their own, without a specific connection to outcome or impact data. Such studies cannot determine causality but, as discussed above, they can be used to address issues critical to policymakers and program administrators, such as the role of the workforce system in career pathway programs. They can provide a detailed picture of program operations, often at a number of locations, and can result in a deeper understanding of the potential and limitations of interventions, as well as generate theories for study. For example, detailed implementation studies could help identify career pathways program components, and combinations of components for future impact research.

A limitation of stand-alone implementation studies, in addition to not being able to answer questions of causality, is that they are unlikely to be statistically representative, and thus generalizing the findings can be problematic. While the rigor of this type of study can be improved through the inclusion of more sites and careful site selection, this will also increase the cost of the study. Overall, however, implementation studies can be conducted at a lower cost than causal studies while yielding important information to the field and identifying areas for future inquiry.

### **6.2.3 Approaches to Addressing Questions about Career Trajectories**

Design options for answering descriptive questions about workers' career trajectories in the labor market are best determined by examining the specific objectives of the planned research inquiry. For example, an analysis of career trajectories using government survey data provides a high level of generalizability because it will provide reliable estimates of trajectories that might represent the entire economy, while data from large firms provides less rigor because it is unlikely to be representative of the population of interest. Studies based on qualitative interviews of workers in a targeted industry may be even less representative because they will likely include smaller numbers of workers.

Because career pathways programs are by design focused on targeted industries, the tradeoff between a large representative survey and a smaller, more focused set of data in terms of generalizability may be merited if researchers are examining very specific questions about the nature of a particular career trajectory in a particular industry, such as trying to better understand barriers to advancement and why previous efforts to addressing those barriers have not succeeded. In such cases qualitative interviews may be the best way to access data rich enough to address the questions of interest. As discussed earlier, one way to obtain both representativeness and insight into a specific career trajectory or industry is to conduct a new survey; that can be very expensive but perhaps is worth considering for a very large industry, such as healthcare, given its share of employment and rate of employment growth (U.S. Bureau of Labor Statistics, 2015). Another option, again as discussed above, is to collect data from large employers in the industries of interest; however, such information is not as helpful when workers transition between employers.

### **6.2.4 Approaches to Addressing Cost Questions**

Cost studies are inextricably bound up in their relationship to outcome and impact data. Cost analysis should use observed study participant and program data wherever possible to measure impacts and outcomes, and are often part of such larger studies, but complete data on all benefits and costs are not always available. If data are not available, imputation techniques and assumptions regarding values

for costs or benefits can be used to complete the analysis. This includes taking benchmarks from external research and developing projections about future earnings based on observed data.

The results of cost-benefit analysis can hinge on a variety of assumptions and projections such as those discussed in Chapter 5. For that reason it is critical that researchers conduct a sensitivity analysis for each assumption to assess how important these assumptions are to the conclusions reached.

### 6.3 Where to Ask: Location Considerations

With tradeoffs associated with the “what” and the “how” addressed, we now turn to issues of “where.” Selection of locations for research is driven by where one can find career pathway programs in the industries and with the program components and populations of particular interest. In selecting sites for future research, DOL should consider establishing site/program selection criteria that will offer the best potential for testing the research questions of interest and allow for collection of high-quality data and rigorous analysis

The type of training provider is a factor in identifying locations that are or should be the subject of future research. The traditional setting for career pathways and related research has been through a range of training providers, most prominently community and technical colleges. These institutions—having played a key role in providing career pathways training to individuals who are seeking to enhance their skills or to enter new fields entirely—are especially well suited to providing career pathways faculty, programs of study, facilities, accelerated completion, and accredited credentials. However, as the career pathway field grows, there may also be other locations that provide strong evaluation opportunities.

More broadly, the following are some criteria that DOL should consider in targeting and/or recruiting locations, sites, or programs for future research. The selected locations should:

- Have a well-developed career pathways program (one that includes many elements of the WIOA definition, for example, and has been operating for some time)
- Offer more than one pathway step of education and training, to sharpen the contrast with traditional job training approaches
- Operate at sufficient scale to evaluate the program (given selected design)
- Demonstrate a local labor market demand for the training, skills, and credentials that the program provides
- Have strong and established partnerships with employers
- Target a set of specific sectors and occupations (e.g., allied health, manufacturing, nursing, data science), including cross-sector occupational clusters (e.g., accounting, information technology, office administration).
- Provide comprehensive supports, such as case management, financial aid and/or Individual Training Accounts, transportation and child care assistance, etc.



- Target one or more specific underrepresented subpopulations (e.g., people of Hispanic ethnicity, youth, people with criminal backgrounds) and/or target occupations and sectors that typically attract underrepresented subpopulations of interest
- Be committed to participating in the evaluation, if commitment to its protocols is needed to implement the research design

### **6.4 Conclusion**

In sum, we are excited about the future of research on career pathways initiatives to answer career pathways research questions critical to policymakers and program administrators in the workforce development field as well as elsewhere. This report has explored a range of research design options available to address different types of questions; considerations such as rigor, cost, and appropriate fit for the questions being asked can help guide DOL in setting its future career pathways research agenda.

## References

- Anderson, T., Kuehn, D., Eyster, L., Barnow, B., & Lerman, R. (2017). *New Evidence on Integrated Career Pathways. Final Impact Report for Accelerating Opportunity*. Washington, DC: The Urban Institute.
- Autor, D. (2015). Skills, education, and the rise of earnings inequality among the “other 99 percent.” Presentation to SNS Seminar, Stockholm, Sweden, August 2015.  
[https://hceconomics.uchicago.edu/sites/default/files/file\\_uploads/autor\\_2015\\_presentatin.pdf](https://hceconomics.uchicago.edu/sites/default/files/file_uploads/autor_2015_presentatin.pdf).
- Barnow, B., & Trutko, J. (2010). “Implementing Efficiency Measures for Employment and Training Programs Final Report.” Washington, DC: U.S. Department of Labor, Employment and Training Administration, Occasional Paper 2010-05.
- Bell, S. H., & Orr, L. (1994). “Is Subsidized Employment Cost Effective for Welfare Recipients? Experimental Evidence from Seven State Demonstrations.” *Journal of Human Resources* 29 (1): 42-61.
- Bell, S. H. & Peck, L. R. (2016). “On the ‘How’ of Social Experiments: Experimental Designs for Getting Inside the Black Box.” *New Directions for Evaluation*, 152: 97-107. DOI: 10.1002/ev.20210
- Berger, M., Black, D., & Smith, J. (2001). “Evaluating Profiling as a Means of Allocating Government Services,” in Michael Lechner and Friedhelm Pfeiffer (eds.), Econometric Evaluation of Active Labour Market Policies, Heidelberg: Physica, 59-84
- Black, D., McKinnish, T., & Sanders, S. (2005). “The Economic Impact of the Coal Boom and Bust.” *Economic Journal* 115(503): 449-476.
- Bloom, H. S., Hill, C. J., & Riccio, J. A. (2003). “Linking Program Implementation and Effectiveness: Lessons from a Pooled Sample of Welfare-To-Work Experiments.” *Journal of Policy Analysis and Management* 22 (4): 551-575.
- Card, D., Kluve, J., & Weber, A. (2010). “Active Labour Market Policy Evaluations: A Meta-Analysis.” *Economic Journal* 120 (548): F452-F477.
- Card, D., & Krueger, A. (1993). “Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania.” Cambridge, MA: National Bureau of Economic Research, Working Paper no. w4509.
- Carnevale, A. P., Jayasundera, T., & Gulish, A. (2016). *America’s Divided Recovery: College Haves and Have-Nots*. Washington, DC: Georgetown University Center on Education and the Workforce.
- Cheng, I., Koralek, R., Robinson, A., Russell, S., & Schwartz, D. (2018). *Career Pathways in Early Care and Education*. (report submitted to U.S. Department of Labor, Chief Evaluation Office).
- Clark, L. (2012). “Google’s Artificial Brain Learns to Find Cat Videos.” *Wired*, June 26.

- Cohen, M. A. (2004). *The Costs of Crime and Justice*. Abingdon, UK; and New York, NY: Routledge.
- Coleman, J. (2017). "Lifelong Learning is Good for Your Health, Your Wallet, and Your Social Life", *Harvard Business Review*, Feb. 7. Available at <https://hbr.org/2017/02/lifelong-learning-is-good-for-your-health-your-wallet-and-your-social-life>
- Collins, L. M., Murphy, S. A., Nair, V. N., & Strecher, V. J. (2005). "A Strategy for Optimizing and Evaluating Behavioral Interventions." *Annals of Behavioral Medicine* 30 (1): 65-73. doi: 10.1207/s15324796abm3001\_8.
- Congressional Budget Office. (2015). *Effective Marginal Tax Rates for Low- and Moderate-Income Workers in 2016*. Washington, DC:
- Dehejia, R. H., & Wahba, S. (2002). "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *The Review of Economics and Statistics* 84 (1): 151-161.
- Dolphin, S., & Schochet, P. Z. (2012). *The Benefits and Costs of the Trade Adjustment Assistance (TAA) Program Under the 2002 Amendments*. Washington, DC: U.S. Department of Labor, Employment and Training Administration. Available at [https://wdr.doleta.gov/research/FullText\\_Documents/ETAOP\\_2013\\_09.pdf](https://wdr.doleta.gov/research/FullText_Documents/ETAOP_2013_09.pdf)
- Elliott, M., & Roder, A. (2017). *Escalating Gains: Project QUEST's Sectoral Strategy Pays Off*. New York, NY: Economic Mobility Corporation.
- Epstein, D., & Klerman, J. A. (2016). "On the 'When' of Social Experiments: The Tension between Program Refinement and Abandonment." In L. R. Peck (Ed.), *Social Experiments in Practice: The What, Why, When, Where, and How of Experimental Design and Analysis*. *New Directions for Evaluation*, 152, 33-45.
- Garcia, J. L, Heckman, J., Ermini Leaf, D., & Prados, M. J. (2017). "Quantifying the Life-cycle Benefits of a Prototypical Early Childhood Program." NBER Working Paper No. 23479.
- Glosser, A., Hamadyk, J., Gardiner, K., & Fishman, M. (2014). *Innovative Strategies for Increasing Self-Sufficiency Career Pathways Profile: Washington's Integrated Basic Education and Skills Training (I-BEST) Program in Three Colleges*. OPRE Report #2014-38, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Hendra, R., & Hamilton, G. (2015). "Improving the Effectiveness of Education and Training Programs for Low-Income Individuals: Building Knowledge from Three Decades of Rigorous Experiments." In *Transforming U.S. Workforce Development Policies for the 21st Century*, edited by Tammy Edwards, Todd Greene, and Carl Van Horn, 411-440. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

- Holl, D. B., Kolovich, L., Bellotti, J., & Paxton, N. (2009). *Evaluation of the Prisoner Re-Entry Initiative*. Washington, DC: Employment and Training Administration, U.S. Department of Labor. Available at [https://wdr.doleta.gov/research/FullText\\_Documents/Evaluation%20of%20the%20Prisoner%20Re-Entry%20Initiative%20-%20Final%20Report.pdf](https://wdr.doleta.gov/research/FullText_Documents/Evaluation%20of%20the%20Prisoner%20Re-Entry%20Initiative%20-%20Final%20Report.pdf)
- Isaacs, J. (2008). *The Costs of Benefit Delivery in the Food Stamp Program: Lessons from a Cross-Program Analysis*. Washington, DC: U.S. Department of Agriculture, Economic Research Service. Available at <http://naldc.nal.usda.gov/download/32783/PDF>
- Jaciw, A. P., (2016). “Applications of a Within-Study Comparison Approach for Evaluating Bias in Generalized Causal Inferences from Comparison Groups Studies.” *Evaluation Review* 40 (3): 241-276. doi: 10.1177/0193841X16664457.
- Jenkins, D., Zeidenberg, M., & Kienzl, G. (2009). Educational Outcomes of I-BEST, Washington State Community and Technical College System’s Integrated Basic Education and Skills Training Program: Findings from a Multivariate Analysis. CCRC Working Paper No. 16. New York, NY: Community College Research Center, Columbia University.
- Kambourov, G., & Manovskii, I. (2013). “A Cautionary Note on Using (March) CPS and PSID Data to Study Worker Mobility.” *Macroeconomic Dynamics* 17(1): 172-194.
- Karoly, L. A., (2008). *Valuing Benefits in Benefit-Cost Studies of Social Programs*. Santa Monica, CA: RAND Corporation. Available at [https://www.rand.org/pubs/technical\\_reports/TR643.readonline.html](https://www.rand.org/pubs/technical_reports/TR643.readonline.html)
- Kelleher, J. D., Mac Namee, B., & D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. Cambridge, MA: MIT Press.
- King, C. T. (2004). “The Effectiveness of Publicly Financed Training in the United States: Implications for WIA and Related Programs.” In *Job Training Policy in the United States*, edited by Christopher J. O’Leary, Robert A. Straits, and Stephen A. Wandner, 57-100. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- King, G., & Nielsen, R. (2016). “Why Propensity Scores Should Not Be Used for Matching.” (Institute for Quantitative Social Science Working Paper). Cambridge, MA: Harvard University. Available at <https://gking.harvard.edu/publications/why-propensity-scores-should-not-be-used-for-matching>
- Klerman, J., & Rolston, H. (2015). “Administrative Data: When Is It Useful for Estimating Impact?” Abt Perspectives blog. Bethesda, MD; Abt Associates. <http://www.abtassociates.com/Perspectives/December-2015/Administrative-Data-When-is-it-Useful-for-Estimati.aspx>
- Laura and John Arnold Foundation. (2015). *Low-Cost RCTs Are a Powerful New Tool for Building Scientific Evidence about “What Works” to Address Major Social Problems*. Available at [http://evidencebasedprograms.org/wp-content/uploads/2016/09/Low-cost\\_RCTs\\_Concept\\_Paper\\_2015\\_12.pdf](http://evidencebasedprograms.org/wp-content/uploads/2016/09/Low-cost_RCTs_Concept_Paper_2015_12.pdf)

- Martinson, K., Copson, E., Gardiner, K., & Kitrosser, D. (2018). *Instituto del Progreso Latino, Carreras en Salud Program: Implementation and Early Impact Report*, OPRE Report # 2018-06, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Martinson, K., & Hamilton, G. (2011). *Providing Earnings Supplements to Encourage and Sustain Employment: Lessons from Research and Practice*. New York, NY: MDRC. Available at [https://www.mdrc.org/sites/default/files/policybrief\\_30.pdf](https://www.mdrc.org/sites/default/files/policybrief_30.pdf)
- Murphy, S. (2005). "An Experimental Design for the Development of Adaptive Treatment Strategies." *Statistics in Medicine* 24 (10): 1455-1481.
- Nisar, H., Klerman, J. A., & Juras., R. (2013). *Designing Training Evaluations: New Estimates of Design Parameters and their Implications for Design*. Bethesda, MD: Abt Associates Inc. Working Paper.
- Office of Management and Budget. (2003). *Circular A-4: Regulatory Analysis*. Washington, DC: Executive Office of the President. <http://www.whitehouse.gov/omb/circulars/>.
- Peck, L. R. (2017) "Yes, You Can Generalize from Experiments!" Abt Associates Inc. Perspectives blogpost, August 31. Available at: <http://www.abtassociates.com/Perspectives/August-2017/Yes,-You-Can-Generalize-from-Experiments!.aspx>
- Peck, L. R. (2015). "Using Impact Evaluation Tools to Unpack the Black Box and Learn What Works." *Journal of MultiDisciplinary Evaluation*, 11 (24): 54-67.
- Peck, L. R., Werner, A., Harvill, E., Litwok, D., Moulton, S., Rulf Fountain, A., & Locke, G. (2018). Health Profession Opportunity Grants (HPOG 1.0) Impact Study Interim Report: Program Implementation and Short-Term Impacts, OPRE Report 2018-16. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Rolston, H., Copson, E., & Gardiner, K. (2017). *Valley Initiative for Development and Advancement: Implementation and Early Impact Report*, OPRE Report #2017-83, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Sarna, M., & Strawn, J. (2018). *Career Pathways Implementation Synthesis*. (report submitted to U.S. Department of Labor, Chief Evaluation Office).
- Schwartz, D., Strawn, J., & Sarna, M. (2018). *Career Pathways Research and Evaluation Synthesis*. (report submitted to U.S. Department of Labor, Chief Evaluation Office).
- Seefeldt, K. (2017). *Progress on Pathways: Findings from Qualitative Interviews with PACE Participants*. OPRE Report #2017-113. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

U.S. Bureau of Labor Statistics, Employment Projections Program. (2015). *Industry employment and output projections to 2024*. Available at <https://www.bls.gov/opub/mlr/2015/article/industry-employment-and-output-projections-to-2024.htm>.

U.S. Department of Education. (2015). *The Evolution and Potential of Career Pathways*. Washington, DC: U.S. Department of Education, Office of Career, Technical, and Adult Education. Available at <http://connectingcredentials.org/wp-content/uploads/2015/05/The-Evolution-and-Potential-of-Career-Pathways.pdf>

U.S. Department of Labor. (2015). Career pathways toolkit: A guide for system development. Washington, DC. Available at [https://www.doleta.gov/usworkforce/pdf/career\\_pathways\\_toolkit.pdf](https://www.doleta.gov/usworkforce/pdf/career_pathways_toolkit.pdf)

Werner, A., Rappaport, C., Stuart, J. B., & Lewis, J. (2013). *Literature Review: Career Pathways Programs* (OPRE Report #2013-24). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Zeidenberg, M., Cho, S., & Jenkins, D. (2010). “Washington State’s Integrated Basic Education and Skills Training Program (I-BEST): New Evidence of Effectiveness.” (CCRC Working Paper No. 20). New York: Columbia University, Teachers College, Community College Research Center.