REPORT

Comparing Job Training Impact Estimates using Survey and Administrative Data

September 28, 2018

Annalisa Mastri
Dana Rotz
Elías S. Hanno

Submitted to:
U.S. Department of Labor
200 Constitution Avenue, NW
Washington, DC, 20210
Project Officer: Janet Javar
Contract Number: DOLQ129633249/DOL-OPS-15-U-00183

Submitted by:
Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Sarah Dolfin
Reference Number: 50174

This report was prepared for the U.S. Department of Labor (DOL), Chief Evaluation Office by Mathematica Policy Research, under contract number DOLQ129633249. 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.

CONTENTS

This page has been left blank for double-sided copying.

TABLES

This page has been left blank for double-sided copying.

FIGURES

This page has been left blank for double-sided copying.

EXECUTIVE SUMMARY

Administrative data can enable researchers to cost-effectively answer a wide range of policy questions. However, the choice of data source—including sources that exist for non-research purposes—can affect the accuracy and precision of a study's earnings impact estimates and policy decisions based on those estimates. To better understand the implications of using different data sources and inform study design decisions, it is therefore important to conduct research on the characteristics, strengths, and weaknesses of different kinds of data.

In this report, we document and explore the strengths and drawbacks of data sources commonly used to produce impact estimates for evaluations of workforce development programs. Specifically, we use information from the Workforce Investment Act (WIA) Adult and Dislocated Worker Programs Gold Standard Evaluation to examine three data sources used to evaluate the impacts of access to services provided by the public workforce system's Adult and Dislocated Worker programs, two of the largest, publicly-funded workforce development programs in the nation. The three sources we consider are: (1) evaluation-administered survey data for the study sample; (2) administrative data from the Administration for Children & Families' National Directory of New Hires; and (3) administrative tax data collected by the Internal Revenue Service, a bureau of the Department of the Treasury. We compare earnings levels and impact estimates based on all three sources to gather insights about the strengths, drawbacks, and tradeoffs inherent in pursuing each data source to answer research questions.

We found that all three data sources were consistent in the main qualitative conclusions about the effectiveness of Adult and Dislocated Worker program services. In addition, the findings across data sources were consistent with previous studies of the differences in impacts between survey and administrative data, which have found that impacts calculated using administrative data tend to be smaller than those calculated using survey data. However, the consistency in the main qualitative conclusions conceals underlying complexities, as several factors seem to move the earnings reported through each data source in different directions at various times during the follow-up period.

The evidence suggests that both surveys and administrative data have strengths and drawbacks when used to assess employment and earnings outcomes for research purposes. Researchers should be cautious about limiting a study to one type of data source and should choose sources based on which is best suited to the policy question or population being studied. For example, administrative data sources do not include some types of employment that are becoming more important with the growth of the gig economy. Moreover, different administrative datasets also have different strengths and weaknesses, such as whether data are available quarterly or annually and the extent to which informal employment and alternative work arrangements are captured. Therefore, if a study population includes people likely to be in informal or alternative labor arrangements, a survey might be more useful than administrative data. Conversely, when reporting on jobs held early in the follow-up period, survey respondents are more likely to forget (or otherwise fail to report) having held some jobs but tend to report having earned more than administrative sources would suggest for each of the jobs they do report. Therefore, for a study with a long follow-up period, administrative data sources that contain the outcomes of interest might be the best option.

Regardless of which type of data a study uses, it is important for researchers and consumers of research to be cognizant of data's strengths and limitations in addressing the research questions of interest, particularly when interpreting impacts derived solely from one data source.

I. INTRODUCTION

Newly available data sources have made it possible to cost-effectively answer a wide range of policy questions. Even for information on earnings, which researchers have long used secondary data to measure, new administrative databases have recently become available for research purposes. For example, since 2015, Federal agencies have made the NDNH more widely available to researchers (Barnow and Greenberg 2015). However, the choice of data source—including sources that exist for non-research purposes—can affect the accuracy and precision of a study's earnings impact estimates and policy decisions based on those estimates. To better understand the implications of using different data sources and inform study design decisions, it is therefore important to conduct research on the characteristics, strengths, and weaknesses of different kinds of data.

In their wide-ranging survey of the literature on how the choice of data sources influenced earnings impact estimates, Barnow and Greenberg (2015) showed that impact estimates based on earnings reported in administrative data tended to be smaller in magnitude and somewhat less likely to be statistically significantly different from zero than those based on earnings measured in survey data. Using data on impact estimates from eight different social experiments, all of which examined earnings outcomes, the authors found that they often could not rule out the possibility that differences in earnings impacts across survey and administrative data sources were large. For example, in the National Job Corps Study (Schochet et al. 2003), impacts on earnings in the 15th quarter after random assignment were more than twice as large when earnings were measured using survey data than when using administrative Unemployment Insurance (UI) data; in the 16th quarter after random assignment, the impacts were seven times as large. Similarly, that study estimated that the impacts of Job Corps on earnings were more than three times higher using survey data than when using administrative Social Security records. This literature has not historically tested differences between impact estimates for statistical significance and it is not generally possible to accurately assess statistical significance based only on published information.[1]

Differences between estimates developed using different data sources could result from differences in the individuals included in the data source, reporting errors, or purposeful omissions. For survey data, differences in response rates can cause survey nonresponse bias. For instance, individuals in the study treatment group, who are eligible to receive employment services, might be more likely to be in longer-term contact with the program and hence more likely to respond to a survey administered by the study team. If those individuals are also more likely to be working, this can introduce bias into the estimated earnings impacts. Reporting errors refer to inaccuracies in the earnings levels recorded, whether that be on an administrative data

[1] Given the expectation that individuals' earnings reports will be at least somewhat correlated across data sources, accurately assessing whether differences between the data sources could reasonably be explained by chance requires accounting for covariance in estimates from different data sources. Significance tests conducted after the fact by comparing individual data sources' estimated impacts and standard errors do not account for this. As we discuss in more detail in describing our methods, this can be difficult or impossible even with access to original data because restrictions are often imposed on what information can be directly merged onto or compared with administrative earnings data sources. These complications, together with the fact most of the literature considers comparing data sources an exploratory addition to work focused on specific program impacts, can explain why formal hypothesis testing was not generally undertaken.

source or a survey, while omissions are inaccuracies that occur when a data source's original purpose differs from the purpose of a research effort. For instance, UI records and administrative tax data sources each cover specific types of income and do not include wages earned in different, specific types of labor. Hence, such data can be expected to understate a given person's total earnings from employment or to suggest those working only in excluded positions are unemployed, which is typically the outcome of interest to researchers. This is an omission from the data sources due to the fact that the goals of the researchers and agencies collecting the administrative data differ. In contrast, surveys often collect respondents' hours worked per week and hourly wage; researchers multiply the two to arrive at weekly earnings. But this approach is subject to inaccuracies in recall about the hours worked and hourly wages (that is, reporting errors). Barnow and Greenberg (2015) concluded that biases that result from reporting errors and omissions seemed to drive more of the difference in impact estimates across data sources than survey nonresponse bias.

Survey data and administrative records are subject to different potential biases, and it is unclear which source provides the most accurate earnings impact estimates when these data are used for research purposes. Survey data rely on respondents truthfully and accurately reporting their employment and earnings histories. Systematic patterns of survey nonresponse—especially differential nonresponse by program participants and non-participants—or misreporting of earnings (either accidentally or purposefully) by respondents can lead to biased estimates (Bollinger and Hirsh 2013; Bound et al. 2000). Likewise, if individuals or firms misreport information to the agencies that create and maintain administrative records, these sources can also produce biased impact estimates when used for research purposes, even while fulfilling the mission for which the data were collected in the first place (Abraham et al. 2013; Blakemore et al. 1996). Furthermore, because administrative data sets are usually developed for purposes other than research, many include information on earnings from only a subset of jobs, reflecting the original purpose of the data collection. For example, the National Directory of New Hires (NDNH), which brings together Federal government employment records and all states' UI records, does not include earnings in private-sector jobs not covered by UI—such as informal work, certain types of farm labor, independent contractor positions, and some other alternative labor arrangements. Administrative tax records include only taxable earnings but should capture all such income unless it is purposefully or mistakenly misreported (for example, to reduce individuals' tax burden). For researchers using these administrative data sources for analyses of total earnings from employment, this will lead to systematic underreporting of earnings levels, but could produce upward- or downward-biased impact estimates.

In this report, we document and explore strengths and drawbacks of data sources commonly used to produce impact estimates for evaluations of workforce development programs using new data and a novel approach to compare estimates across data sources. Specifically, we use information from the Workforce Investment Act (WIA) Adult and Dislocated Worker Programs Gold Standard Evaluation to examine three data sources used to evaluate the impacts of access to services provided by the public workforce system's Adult and Dislocated Worker programs,[2] two of the largest publicly funded workforce development programs in the nation: (1)

[2] Our findings are based on individuals who were enrolled in the study between 2011 and 2013, when the Adult and Dislocated Worker programs operated under the Workforce Investment Act of 1998. In 2014, the programs were reauthorized by the Workforce Innovation and Opportunity Act.

evaluation-administered survey data; (2) administrative data from the Administration for Children & Families' NDNH; and (3) administrative tax data collected by the Internal Revenue Service (IRS), a bureau of the Department of the Treasury. The expanding availability of the NDNH as an administrative data source with nearly nationwide coverage has garnered substantial interest as a more convenient, comprehensive, and cost-effective source of UI wage records than compiling them state by state (Barnow and Greenberg 2015). Administrative tax data are another source of income data that provide nationwide coverage, including covering some jobs that the UI records underpinning the NDNH do not report.

By comparing earnings levels and impact estimates based on all three sources, we shed light on the differences between the data sources, and gathered insights about the tradeoffs inherent in pursuing each. In general, we found that administrative and survey data each have shortcomings, but they complement each other well. Survey data can contain a richer set of outcomes, over a more relevant time period, but the quality varies over time, with accuracy declining substantially the farther back respondents are asked to recall the details of their jobs. Meanwhile, the NDNH and administrative tax data (by their design) fail to capture earnings from certain types of employment, which can lead to underestimates of earnings. These two types of shortcoming, and their intersection over time, explain the patterns observed in the earnings impact estimates. Generally, our findings concur with those of Barnow and Greenberg (2015), that reporting errors seem to drive much more of the difference in estimated earnings impacts than survey nonresponse bias.

In Section II, we describe in more detail the WIA Adult and Dislocated Worker Programs Gold Standard Evaluation (WIA Gold Standard Evaluation), the three particular data sources we examined, and the advantages and disadvantages to each. Section III discusses our methods for estimating impacts across the data sources and the results of that analysis. Section IV presents our methods for investigating the differences in means and impacts across two of the data sources and the results of that analysis. Section V presents our conclusions.

This page has been left blank for double-sided copying.

II. THREE SOURCES OF DATA ON EARNINGS AND EMPLOYMENT

This study drew on three data sources measuring earnings and employment covering study participants in the WIA Gold Standard Evaluation: the WIA Gold Standard Evaluation surveys, the NDNH, and administrative tax data (analyzed separately by Manoli and Patel 2018).[3] The first section of the chapter provides a brief overview of the WIA Gold Standard Evaluation. The next three sections describe the three data sources, and the final section presents the relative merits of each.

A. Background on the WIA Gold Standard Evaluation

The WIA Gold Standard Evaluation estimated the impacts and cost-effectiveness of intensive and training services provided through the WIA Adult and Dislocated Worker programs and examined how sites implemented the programs nationwide. In the present study, we focus on the impact component of the WIA Gold Standard Evaluation.

The study team designed the impact evaluation to produce nationally representative estimates of the effects of the Adult and Dislocated Worker programs. Two key characteristics of the study design enabled this: (1) random selection of 28 local areas for the study and (2) random assignment to study groups of nearly 36,000 customers served by the Adult and Dislocated Worker programs within those areas (Mastri et al. 2015).

Depending on their needs and eligibility, customers of the Adult and Dislocated Worker programs could receive services through three tiers: (1) core services, which either were self-directed (accessed from resource rooms located at the American Job Centers or via the Internet) or required a modest amount of staff assistance; (2) intensive services, which generally required more staff assistance than core services; and (3) training (Figure II.1). WIA required that local areas provide core, intensive, and training services sequentially. That is, customers had to receive at least one core service to receive an intensive service and at least one intensive service to receive funding for training.

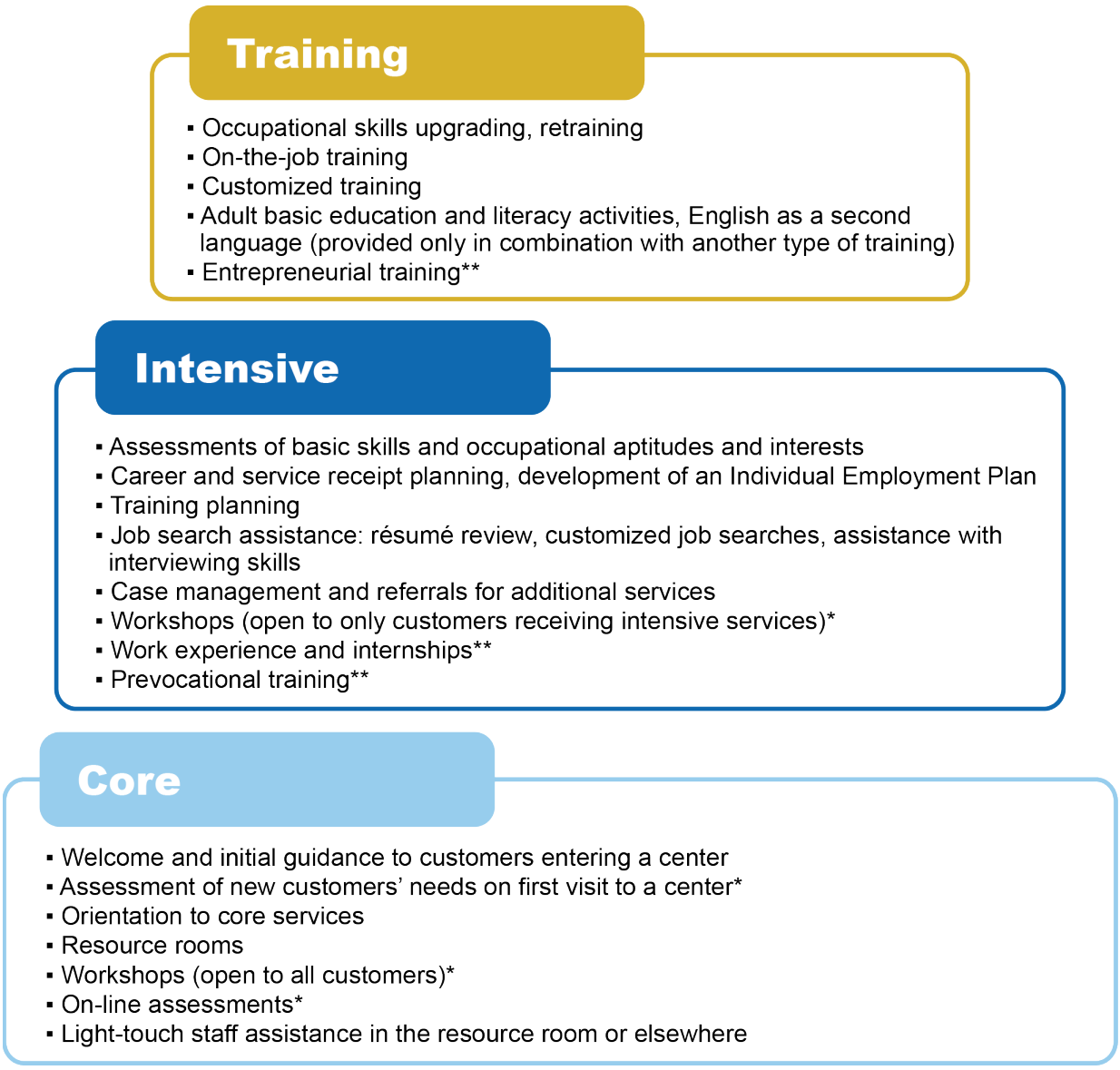
Staff randomly assigned eligible and consenting customers to one of three study groups:

1. **Full-WIA group.** Customers in this group could receive any core, intensive, or training services for which they were eligible, just as they would in the absence of the evaluation. The study did not require offering training to everyone in this group, nor that everyone offered training actually enroll in a training program. The study team randomly assigned most study participants (31,304 customers or 88 percent of those randomly assigned) to the full-WIA group.

[3] We note that this paper is relevant for tax policy. In particular, this paper is derived from a project at the Office of Tax Analysis titled "The Effects of Employment and Earnings on Tax Filing and Tax Liability: Evidence on Short-term and Long-term Effects Using Administrative Tax Data." Since forecasting tax filing and tax liabilities are central components of tax policy, this project aims to understand how changes in employment and earnings affect changes in tax filing and tax liabilities. This project exploits exogenous variation in employment and earnings to estimate the causal effects of changes in employment and earnings on tax filing and tax liabilities. While this report refers to research based on administrative tax data, this report does not represent any views or opinions from Manoli and Patel (2018) or any views or opinions from the United States Treasury or Internal Revenue Service.

2. **Core-and-intensive group.** Customers in this group could receive any core or intensive services for which they were eligible but they *could not receive training* services funded by the Adult or Dislocated Worker programs. No customer was required to receive intensive services. The study team randomly assigned 2,181 customers (6 percent of those randomly assigned) to the core-and-intensive group.
3. **Core group.** Customers in this group could receive *only core services* and no intensive or training services funded by the Adult and Dislocated Worker programs. The study team randomly assigned 2,180 customers (6 percent of those randomly assigned) to the core group.

Figure II.1. Types of services offered by local areas in the study



Source: D'Amico et al. (2015).

*In some, but not all, local areas.

**Rarely offered, and offered to only a small number of customers.

Customers' access to services was restricted according to their study group for 15 months after random assignment. After this period, customers could receive any services offered by the programs for which they were eligible. The main sources of service receipt and outcomes data were follow-up surveys conducted 15 and 30 months after random assignment and administrative data from the NDNH. McConnell et al. (2016) and Fortson et al. (2017) present the results of the impact analyses.

B. WIA Gold Standard Evaluation surveys

The research team for the WIA Gold Standard Evaluation conducted two telephone surveys, 15 and 30 months after they randomly assigned customers, to collect data on customers' service receipt and outcomes, focusing especially on employment and earnings. The research team attempted to contact all members of the core-and-intensive and core groups. Because the full-WIA group contained a disproportionate share of study participants, the team attempted to collect data from about 2,000 members of this larger study group using random sampling.

Both surveys collected information on a variety of measures. The 15-month survey asked for information about customers' service receipt, participation in training, and employment and earnings since random assignment. It also requested details on the characteristics of the jobs workers held, including the industry, occupation, hourly wage, benefits, and average hours per week associated with each job, which can provide more detailed and nuanced information on respondents' experience in the labor market. For respondents to the 15-month survey, the 30-month survey asked for information on these same outcomes starting on the date of their most recent interview. For participants who did not respond to the 15-month survey, the 30-month survey asked for information covering the entire period after random assignment. Thus, the research team could construct a 30-month timeline of employment for all respondents to the 30-month survey.

For this report, we used information from one or both follow-up surveys to construct several measures of earnings and employment for the 4,777 respondents to the 30-month survey (response rate of 77 percent).[4] For comparability with the NDNH, we constructed quarterly measures of earnings and employment for nine calendar quarters after random assignment and, for comparability with administrative tax data, we constructed a measure of annual earnings for the first full calendar year after random assignment. As in the main impact reports, we constructed weekly earnings measures by multiplying respondents' reported hours worked per week and average hourly wage for each job.[5] We then aggregated and combined these with dates of employment in each job to produce measures of quarterly and annual earnings. Quarterly earnings data for the nine calendar quarters and the first full calendar year after random assignment were available for 98 percent or more of survey respondents. However, because of the timing of the surveys' administration relative to study participants' dates of random assignment (30 months after random assignment), we could construct a measure of annual earnings for the second full calendar year following random assignment only for 44 percent of

[4] Response rates did not differ significantly across the three study groups. All weighted analyses account for survey nonresponse, as described in Rotz et al. (2017).

[5] When survey respondents provided salaries on a monthly, weekly, annual, or other basis, we used reports of weeks worked per year and hours worked per week to estimate the hourly wage.

survey respondents. For example, an individual randomly assigned in March would not start their first full calendar year after random assignment until 10 months after random assignment and their second full calendar year would start 22 months after random assignment. Therefore, the 30-month follow-up period would not include the entirety of the second calendar year for this individual.

C. National Directory of New Hires

The NDNH, an administrative database, contains wage information collected by state unemployment agencies, supplemented with wage information for Federal employees, and submitted to the Office of Child Support Enforcement of the Administration for Children and Families, U.S. Department of Health and Human Services (U.S. Department of Health and Human Services 2014). The database contains information on individuals' quarterly earnings; UI benefits received; and start dates of any new jobs by calendar quarter (that is, January to March, April to June, July to September, or October to December of a given year). Notably, unlike the UI wage records used by many previous studies to analyze the impact of workforce programs (for example, Heinrich et al. 2008, 2013; Andersson et al. 2013), the NDNH contains information from all states and for Federal government employees, enabling researchers to construct measures of earnings and employment that include data from jobs worked in a state other than the one where an individual resides.

For this report, our analysis of the NDNH focused on measures of earnings and employment for the nine calendar quarters following random assignment. These measures were constructed using NDNH data from 2012 to 2015 and were available for 98 percent of study participants, including survey nonrespondents and participants in the full-WIA group who were not sampled for the survey.[6] For comparability with administrative tax data, we aggregated NDNH quarterly data to the annual level and constructed measures of earnings and employment for the first two calendar years following participants' random assignment. A subset of study participants also had sufficient data to construct an annual earnings measure for the third calendar year after random assignment (86 percent of the sample).

D. Administrative tax data

The IRS maintains administrative tax data, which contain information from a variety of tax forms, including tax units' 1040, W-2, and 1099-MISC forms. A tax unit refers to the individual or set of individuals listed on a filed federal income tax return to define the unit on which the federal personal income tax is imposed. In the case of a single individual, the tax unit would correspond to that individual. In the case of a married couple with dependent children, the tax unit would refer to the couple and children. The information in these forms enables researchers to construct a variety of measures, including earnings for each individual in the tax unit (based on W-2 and 1099-MISC forms), disability status (based on the 1099-SSA), educational attendance (based on the 1098-T), and savings decisions (from W-2, 1099-B, and other forms).

[6] We excluded 577 members of the WIA Gold Standard Evaluation study sample (1.7 percent) because the customers provided invalid combinations of name and Social Security number. We also excluded 79 members of the study sample (0.2 percent) who were randomly assigned very late in the random assignment period because the NDNH did not include information on them for the full 36-month period used in the WIA Gold Standard Evaluation's analysis of NDNH data.

Because employers and other third parties also report information to the IRS, the database contains information on both tax filers and nonfilers (see Manoli et al. 2018 for details).

For this report, the analysis included administrative tax data from W-2 and 1099-MISC forms for the calendar years 2012 to 2015. (This data source was not available for the evaluation's main impact analyses.) Using names and Social Security numbers collected by the WIA Gold Standard Evaluation team, Manoli and Patel (2018) matched 99 percent of study participants to this administrative data source. Match rates did not differ significantly across study groups. For each study participant, they constructed a measure of earnings and employment in each calendar year from 2012 to 2015 based on earnings reported on W-2 forms. This enabled the measurement of earnings and employment for almost all study participants for the first two calendar years following random assignment; a subset of study participants also had sufficient data to construct a measure of earnings for the third calendar year after random assignment.

Manoli and Patel (2018) also constructed and analyzed an expanded measure of earnings drawing data from both W-2 and 1099-MISC forms. This measure contains information on payments received as an independent contractor, which the NDNH does not include. Like the NDNH, neither W-2 nor 1099-MISC forms include income from self-employment.

E. The merits of each data source for research purposes

Each of the three data sources used in this study presented unique advantages and disadvantages—that often vary depending on the specific research questions, populations studied, timeline, and study context—for research examining individuals' earnings and employment. These comparisons are made from the perspective of a researcher or funder using the data to conduct research, and not that of the agency that collected or commissioned the collection of the data for its specific purposes or requirements (such as tax collection or child support enforcement).

Cost and convenience. For the purpose of conducting research, survey data are almost always substantially more expensive to acquire than administrative data because surveys require designing a questionnaire, employing staff to administer it to the study sample, and locating sample members who might relocate frequently. This is especially true for large studies and those designed to follow sample members for many years. Administrative data also have some costs when used for research, in terms of financial resources and time and requirements for access. For instance, some administrative data, including administrative tax data, require accessing the data in one of a few secured locations nationwide, as a means of protecting the privacy and safety of personal information. Other administrative data, including the NDNH, require that the agency in control of the data prepares a data set for researchers that contains only the specific variables and records needed for the purposes of the research project. Coordinating between the research team and the agency can involve much more preparation, review, and quality control on the part of both researchers and agency staff than would typically be necessary if the study team accessed the data directly.

Outcomes. Each data source considered contained a different set of outcomes. Because research teams customize surveys to their needs, the richness and detail of the information the surveys generate can be much greater than that available from administrative sources, which are

collected for non-research purposes. Typically, this means the data provide additional outcomes of interest as well as greater opportunity to unpack estimated impacts. In addition—and crucially for many impact evaluations—surveys can collect information on the services received by members of all study groups from all sources.

In contrast to study-designed surveys, when using administrative data sources, researchers are constrained by the types of data already collected for their original purposes. For example, as summarized in Table II.1, the WIA Gold Standard Evaluation surveys collected information on outcomes, background characteristics, and other variables not available in the NDNH and administrative tax data. The outcomes available from different administrative data sources also vary depending on the data sources' original use. For example, NDNH data are based on UI wage records and thus contain earnings for each job an individual had (and that the employer reported to UI), as well as UI benefits. Administrative tax data are based on tax filings and thus contain not only earned income through wage or salary labor, but also income from independent contracting and other kinds of earnings captured through 1099-MISC forms.

Table II.1. Outcomes available for the WIA Gold Standard Evaluation, by data source

	WIA Gold Standard Evaluation surveys	NDNH	Administrative tax data
Coverage period	30 months after random assignment, with exact dates varying by individual	2012–2015	2012–2015
Periodicity of earnings and employment records	Weekly	Quarterly	Annual
Number of full calendar years (January to December) after random assignment in which data are available for all participants in database	1	1	2
Number of full calendar years (January to December) after random assignment in which data are available for at least some participants in database[a]	2	3[a]	3[a]
Calendar quarters (January to March, April to June, July to September, or October to December) for which earnings and employment information is available	9	12	None (annual data only)
Other information collected	Various, including services received, total household income, and insurance	UI benefits, new hires	Various, including total tax unit income, savings measures, and disability status
Sample size	**4,699**	**33,773**	**34,250**

[a] A fourth year of data was available for less than 5 percent of study participants.

Timing. Surveys can capture information on finer-grained time periods than those captured in administrative data. The surveys from the WIA Gold Standard Evaluation collected employment and earnings data based on a detailed job history module asking individuals to report the start date, end date, wage rates, and typical hours worked for each job they held during the follow-up period; the research team used this information to estimate weekly earnings and

then aggregated these measures to a quarterly basis. This allowed the team to estimate earnings for ten 3-month periods, beginning after random assignment, regardless of when random assignment occurred.

Administrative data are also limited by the existing schedule and frequency with which they are collected—quarterly in the case of NDNH and annually for administrative tax data—which does not align with study participants' program enrollment dates.[7] However, administrative data sets have an advantage over surveys in that that they continue to be updated after the survey fielding period ends, providing for longer potential follow-up periods.

Sample. Neither survey, NDNH, nor administrative tax data can be used to access information on all workers that would be ideal for a research study, but the two administrative data sources come close (although, as discussed in the following section, these data may not include all *jobs* held by workers). Surveys contain information on only those sample members who researchers can locate and who agree to participate in the survey. The nonresponse bias this induces has long been well-understood (for example, Madow et al.1983a, 1983b) and is largely avoided by administrative data sources.[8] The NDNH and administrative tax data cover nearly all sample members engaged in reportable work; only those sample members who do not provide the study team with a valid Social Security number and matching name are not included. This implies that if these sources lack records for a study participant with a valid name and Social Security number combination, researchers can conclude the individual had no reportable earnings.

Accuracy of employment information. Each data source has slightly different employment coverage, with notable implications for research. Because survey respondents self-report their jobs, survey data should theoretically cover all types of employment; however, intentional misreporting and—especially when survey respondents are asked to report over long periods of time—recall errors are possible (Bound et al. 2000). Like most surveys, the WIA Gold Standard Evaluation surveys were likely vulnerable to recall error. In particular, the 30-month survey asked customers who did not respond to the 15-month survey to recall their employment history for the full 30-month follow-up period. It is possible that these respondents inadvertently failed to report jobs, especially those held early in the study period.

Employment information available through administrative data sources reflects the administrative use of the data, which might or might not match the measure desired by a researcher. For example, the NDNH is used for enforcement of child support orders, to determine when noncustodial parents have entered into employment and to capture any new

[7] The length of time for which follow-up data was available and when that period started depended on when random assignment occurred. For example, an individual randomly assigned in February 2012 would have post-random assignment data by calendar quarter starting in the second quarter of 2012 and post-random assignment data by calendar year starting in 2013.

[8] However, note that nonresponse bias can be minor if overall and differential rates of nonresponse are low (What Works Clearinghouse 2013, 2014). Nonresponse weights can further mitigate any resultant biases by making the sample of survey respondents more similar to the overall study population of interest (Groves and Peytcheva 2008). For the WIA Gold Standard Evaluation, nonresponse bias was minimal because of the low nonresponse rates, which were similar across the three study groups (between 21 and 24 percent), and use of nonresponse weights.

employment covered by UI in the formal sector. They provide a less accurate measure of employment of any sort, a measure typically of greatest interest to employment researchers. Likewise, the administrative tax data provide an exceptionally accurate measure of employment in jobs generating taxable income reported on W-2 and 1099-MISC forms. However, this measure of employment is narrower than employment of any sort. For example, some forms of self-employment are reported using other tax forms.

Several factors could compromise the accuracy (to researchers) of employment information in the NDNH. While the directory covers almost all individuals who engage in reportable work, it does not provide information on all jobs held by covered individuals because the UI records on which it is based do not include certain types of jobs, such as farm labor or self-employment. The UI wage records that compose the NDNH also do not include many informal and under-the-table jobs, nor are positions that employers rightly or wrongly categorize as independent contracting included; evidence suggests that employers sometimes categorize employees as independent contractors to reduce employers' tax burden (Abraham et al. 2013; Blakemore et al. 1996). Past estimates of the proportion of jobs not reported in UI wage records range from 9 to 14 percent (Blakemore et al. 1996; Hotz and Scholz 2002; Wallace and Haveman 2007). More recent studies of one type of non-reportable job, independent contractor arrangements, find independent contracting accounts for 7 to 15 percent of employment (depending on the data source used), and that this proportion may change significantly over time based on how it is measured (Katz and Krueger 2016; Bureau of Labor Statistics 2018; Abraham et al. 2018). This may not be problematic from the standpoint of the NDNH's original purpose, enforcing child support orders, but is a concern for researchers.

The administrative tax data have similar, although more muted, challenges when used for research purposes. This data source should contain information on earnings in all jobs reported by either employees or employers through W-2 forms. The use of two different information sources would suggest nonreporting was a lesser problem for this data source than for the NDNH. In addition, the expanded measure of earnings for the administrative tax data included earnings received as an independent contractor. However, underreporting of self-employment income was likely still an issue in these data.

Finally, note that employment may be measured differently in survey and administrative data, potentially leading to differences in the estimates across data sources. For example, WIA Gold Standard Evaluation's the survey data counted a person as employed if the respondent held one of his or her reported jobs during the period of interest. In the administrative data, a person was employed if he or she had positive earnings during the period of interest. In the survey and NDNH data, we could also measure the number of jobs individuals held, based on the number of jobs reported and number of jobs with positive earnings, respectively.

Accuracy of earnings information. Even when employment in a specific job is reported accurately, earnings from that job might not be. Previous studies have concluded that survey respondents tend to overreport earnings, perhaps as a result of the way in which researchers collect the information. In their comparison of survey and administrative data, Schochet et al. (2003) concluded that overreporting of weekly hours worked—which the study used to compute weekly earnings—drove much of the difference between data sources (and not overreporting of average hourly wage). Smith (1997) found that asking survey respondents about their earnings in

individual jobs and aggregating that information to an annual total (as Schochet et al. 2003 did) led to higher annual earnings than asking respondents directly for total annual earnings, although this could also stem from respondents underestimating annual earnings. Because the WIA Gold Standard Evaluation took an approach similar to Schochet et al. (2003) in computing weekly earnings from weekly hours worked multiplied by hourly wage rate, this could be a concern for the evaluation's survey data as well. See Barnow and Greenberg (2015) for an extensive review of this literature.

Turning to administrative data, employers reporting wages to government agencies have an incentive to understate earnings for a given job to avoid paying payroll taxes (Blakemore et al. 1996). Despite this, W-2s and other third-party-reported documents have been found to be highly accurate for their intended purpose (identifying taxable wages). For example, in audit studies, third-party reported income is rarely changed. Issues related to misreporting of self-employment income in the administrative tax data are recognized, and there are efforts to improve the accuracy of reporting.

This page has been left blank for double-sided copying.

III. OUTCOMES AND IMPACTS FROM THREE DATA SOURCES

In this chapter, we explore the extent of the differences in estimates of earnings and employment across data sources. Section A describes the impact estimation framework we used, following Fortson et al. (2017). Section B presents our estimates of the average annual earnings and employment of customers in our study sample using information from the survey, NDNH, and administrative tax data; and Section C describes the impacts estimated for these annual measures.

A. Estimating impacts of services funded by the Adult and Dislocated Worker programs using the three data sets

As discussed in Chapter II, Section A, there were three tiers of services provided by the Adult and Dislocated Worker programs under WIA: core, intensive, and training services. The WIA Gold Standard Evaluation randomly assigned customers to one of three groups: (1) the core group, who had access only to core services; (2) the core-and-intensive group, who had access to both core and intensive services, but not training funded by the Adult and Dislocated Worker programs; and (3) the full-WIA group, who had access to all program services, including core, intensive, and training services.

We, along with Manoli and Patel (2018) used a simple approach to estimate the impacts of Adult and Dislocated Worker program services in each of the three data sets. To determine the effect of providing training services, we compared the average outcomes of the full-WIA and core-and-intensive groups. To determine the effect of providing intensive services, we compared the average outcomes of the core-and-intensive and core groups. To determine the effect of providing both training and intensive services, we compared the average outcomes of the full-WIA and core-only groups. This is the same approach used to estimate impacts with survey and NDNH data for the WIA Gold Standard Evaluation, but estimates for the survey and NDNH data might differ slightly from those in Fortson et al. (2017) because we made some small adjustments to the analytic sample and the outcomes to simplify comparisons between the data sets.[9]

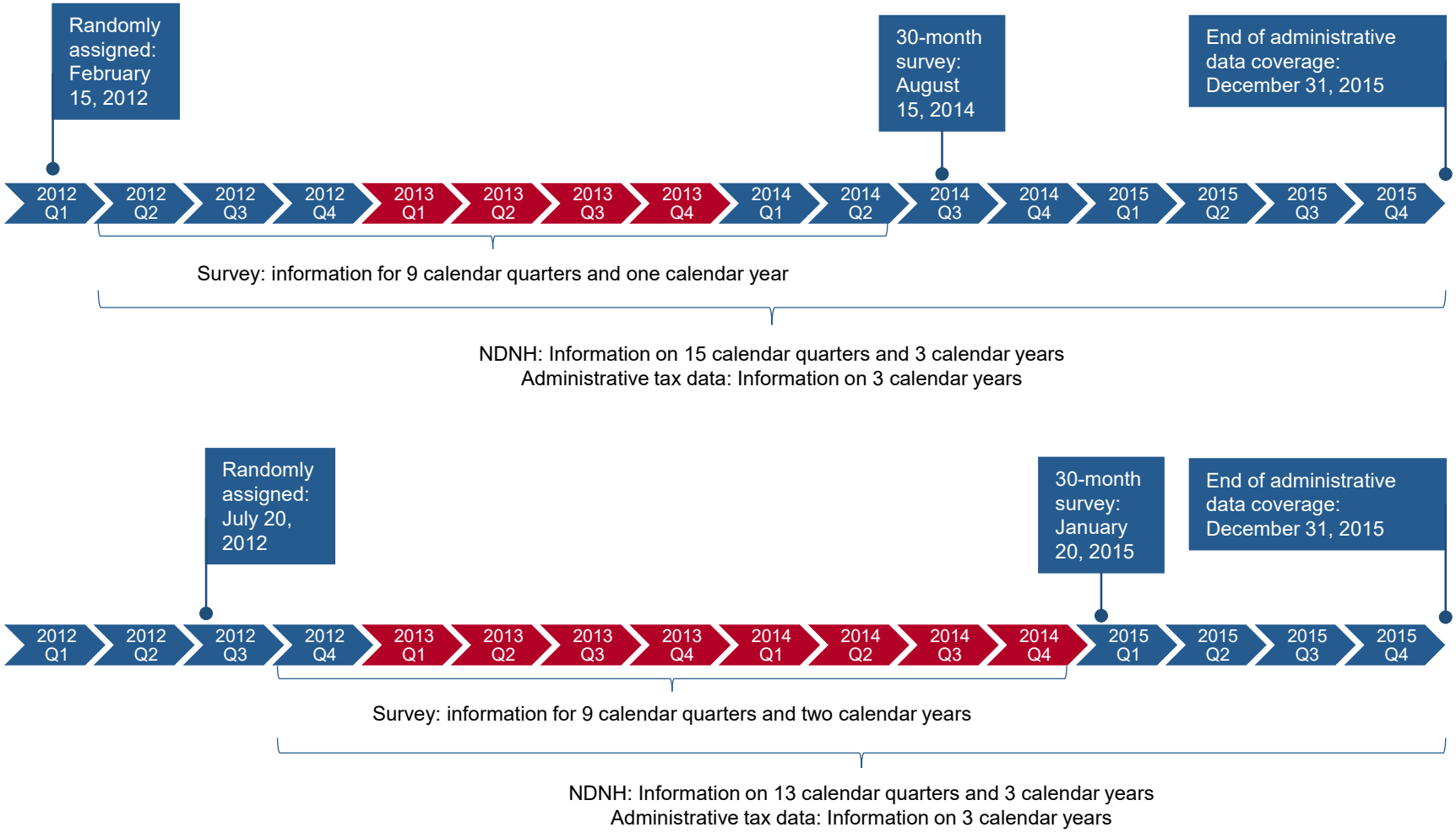
As in the approach used in Fortson et al. (2017), we refined the estimated means and impacts by doing the following:

- Using logical and regression imputation so that missing data did not bias our results[10]

[9] The WIA Gold Standard Evaluation (Fortson et al. 2017) aggregated survey data into quarters based on the number of weeks following random assignment and analyzed NDNH data by calendar quarters, which start at the beginning of the first January, April, June, or September after random assignment. In contrast, we use calendar quarters for both data sources. Appendix Table A.1 compares estimates of earnings from the survey data generated using the two approaches. As one might expect, the differences are large early, when earnings are growing rapidly. In addition, the results in Fortson et al. (2017) might differ from those presented here because we omitted a few individuals with missing NDNH data (1.9 percent of the full sample) from the survey data analytic sample (see footnote 6 for further details) and a few survey respondents with missing earnings information on the survey (2 percent of survey respondents) from the NDNH analytic sample.

[10] Rotz et al. (2017) describes the imputation approach in detail.

Figure III.1. Illustration of the data available for two individuals randomly assigned for the WIA Gold Standard Evaluation on different dates



- Adjusting the impact estimates and standard errors to account for the procedures used to randomly select local areas for the evaluation[11]
- Weighting our estimates for some analyses, so that the impacts generalized to the population of interest

For further details about the estimation approach and regression specifications, see Rotz et al. (2017).

As a result of the variation in coverage, the different data sets contain earnings and employment for different time periods (Figure III.1). In comparing outcomes across all three data sets, we have information on one full calendar year following random assignment for all participants included in each data source; these data can be weighted to produce nationally representative estimates of means and impacts. In addition, the two administrative data sources contain information for all participants for two years after random assignment. Therefore, in the analyses that follow, we produce weighted impact estimates for the first calendar year after random assignment for all three data sources, and for the second and third calendar years after random assignment for the NDNH and administrative tax data.[12] In addition to the comparisons across all three data sources using annual data, with the survey and NDNH data we can analyze impacts and outcomes on a quarterly basis for nine calendar quarters after random assignment.

B. Comparing employment rates and average earnings using annual data

When we weighted the data to be nationally representative, **employment rates were typically similar across data sources when measured by calendar year after random assignment**, ranging from 80 to 84 percent in the first year after random assignment and from 82 to 85 percent in the second year after random assignment (Figure III.2).[13]

Weighted average annual earnings were higher in the survey data than in either administrative source for the first year after random assignment (Figure III.3). In Year 1, weighted survey-reported earnings were $15,511, which is about 14 percent more than the NDNH earnings and about 8 percent more than administrative tax data from W-2s alone. The relatively high survey-reported earnings are consistent with expectations, and could reflect

[11] We estimated impacts using ordinary least squares. Because our primary sample unit for the evaluation was a local area and sampling was stratified by U.S. Department of Labor region, we included region fixed effects in the regression and clustered standard errors at the local area level. In addition, we employed a finite population correction for variance estimation based on an estimate of the share of the population of WIA customers over the follow-up period who were in our sample.

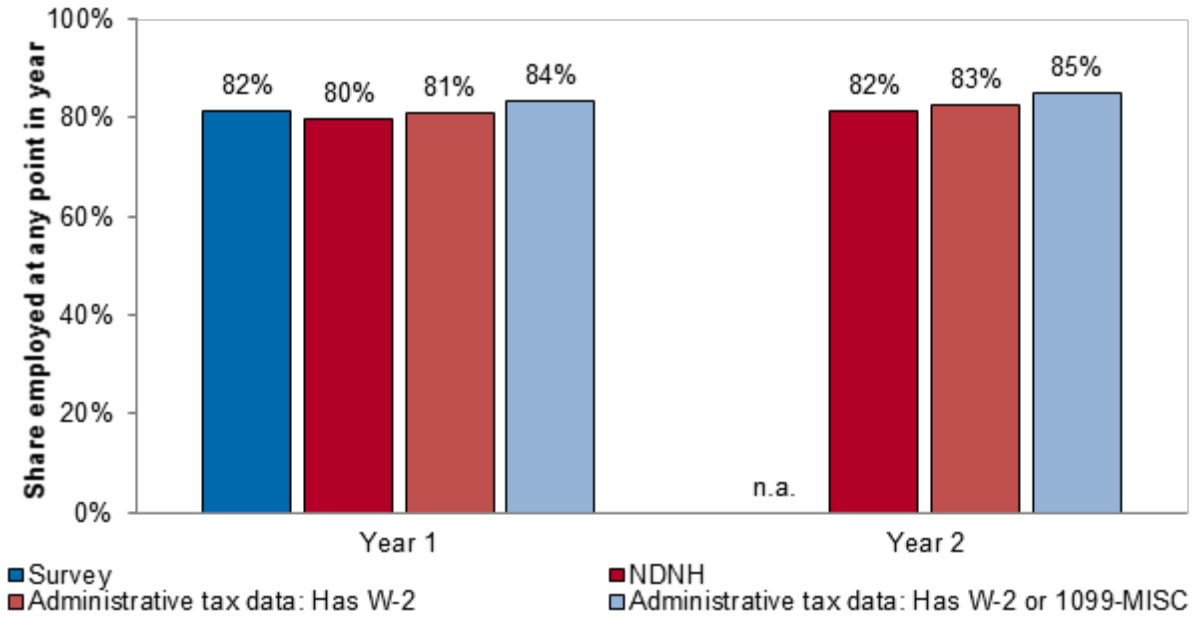
[12] At least some study participants in each data set also have information on earnings and employment for a second year following random assignment, and the administrative data contain information for a subset of participants for three years following random assignment. Therefore, we can use data for the second year after random assignment in the survey and the third year after random assignment in the administrative sources to compare results across sources, but these data cannot be weighted to produce nationally representative estimates. The unweighted analysis enables us to compare results within our sample alone. These results are cited in the text when they differ from the weighted results; the full set of unweighted results is presented in Appendix A.

[13] See Appendix Table A.2 for the corresponding unweighted employment rates, which exhibited similar patterns.

survey overreporting of earnings (which makes survey-reported earnings higher than actual earnings) and underreporting of self-employment income to administrative sources (which makes earnings reported to administrative data lower than actual earnings).That earnings are reportedly higher in the administrative tax data than in NDNH in both years is also consistent with expectations about administrative tax data providing better coverage than the NDNH of certain types of employment, such as some farm labor.

Adding information from 1099-MISC forms suggests that independent contracting is an important source of income. On average, 1099-MISC forms included about $850 of earnings in the first year after random assignment and $900 in the second year after random assignment. When this income is included in measures of earnings, the survey and administrative tax data exhibit similar weighted earnings levels. Therefore, even though the level of income generated by combining the W-2 and 1099-MISC income is similar to that reported in the survey, survey overreporting of earnings and underreporting of self-employment data to administrative sources is still likely taking place.[14]

Figure III.2. Annual employment rates: Data weighted to be nationally representative



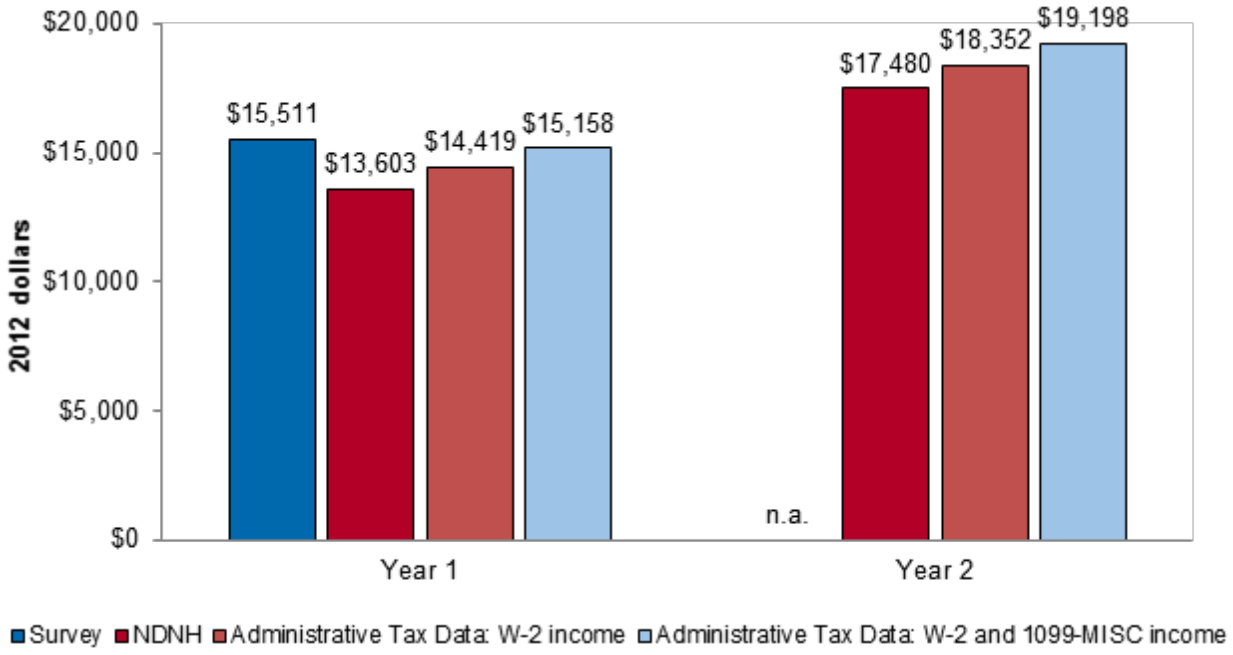
Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys, Manoli and Patel (2018), and National Directory of New Hires.

Note: Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For detailed estimates, see Appendix Table A.2.

n.a. = Not available.

[14] See Appendix Table A.3 for unweighted earnings, which exhibited a similar pattern.

Figure III.3. Annual earnings: Data weighted to be nationally representative



Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys, Manoli and Patel (2018), and National Directory of New Hires.

Note: Estimates include zeroes for those who were not employed in the corresponding time period.[15] Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For detailed estimates, see Appendix Table A.3.

n.a. = Not available.

C. Comparing impacts on earnings and employment rates using annual data

Across all three data sources, the results indicated no statistically significant impacts on employment, with one exception: as measured in the NDNH, the employment rate for study participants in the full-WIA group, which could be offered training and intensive services, was three percentage points lower than the employment rate for those who had access only to core services (Table III.1). This would be expected, as the customers who were eligible for training participated in their training programs (and were not employed simultaneously).

There was no consistent pattern in the relative magnitudes of the employment impact estimates for different data sources. In some comparisons the point estimates were smaller in the administrative data, but in others they were not. Using unweighted data—for which it is possible to examine an additional year using a subset of the study sample—the results again

[15] We do not estimate earnings outcomes conditional on employment (or any earnings) because whether customers are in this group or not is driven in part by WIA Adult and Dislocated Worker Programs services, so such estimates would not be valid estimates of the programs' impacts.

indicated few statistically significant impacts on employment and no consistent patterns in the relative magnitudes of the point estimates based on the data sources (Appendix Table A.4).

Table III.1. Annual impacts on employment: Data weighted to be nationally representative

			Administrative tax data	
	Survey	NDNH	W-2	W-2 and 1099-MISC
Impacts of WIA-funded training				
Employment: (percentage point impact)				
Year 1	-1.8	-1.2	-3.5	-2.6
	(0.313)	(0.487)	(0.230)	(0.341)
Year 2	n.a.	-1.6	-3.8	-3.8
		(0.242)	(0.094)	(0.055)
Impacts of WIA-funded intensive services				
Employment: (percentage point impact)				
Year 1	2.5	-1.7	0.8	0.3
	(0.356)	(0.342)	(0.725)	(0.874)
Year 2	n.a.	1.5	4.2	3.3
		(0.554)	(0.258)	(0.201)
Impacts of WIA-funded training and intensive services				
Employment: (percentage point impact)				
Year 1	0.7	-3.0*	-2.7	-2.3
	(0.667)	(0.033)	(0.057)	(0.106)
Year 2	n.a.	-0.1	0.4	-0.5
		(0.943)	(0.809)	(0.597)
Sample size	**4,699**	**33,773**	**34,250**	**34,250**

Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys Manoli and Patel (2018), and National Directory of New Hires.

Note: Estimated impacts are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. Reported *p*-values for impacts are in parentheses and are based on two-tailed t-tests.

n.a. = Not available.

*Significantly different from zero at the .05 level.

Across all three data sources, the impacts were statistically significant on earnings for two study groups at either the 5 or 10 percent significance level (Table III.2):

- **Access to WIA-funded intensive services increased earnings in the first year after random assignment and, even more consistently across data sources, in the second year after random assignment.** The statistically significant point estimates in Year 1 estimated using survey data and administrative tax data from both W-2 and 1099-MISC forms were very similar in magnitude ($2,710 and $2,544, respectively). In Year 2, for which we could conduct weighted analyses only using the administrative sources, all three estimates of the impact of WIA-funded intensive services on earnings were statistically significant.

- There were marginally statistically significant impacts of access to the combination of WIA-funded training and intensive services. Although only one point estimate—of $1,644 in Year 1, according to survey data—was statistically significant at the five percent level, five of the six point estimates from administrative data were also statistically significant, albeit only at the 10 percent level. In Year 2, the estimates from the administrative data sources are significant at the 10 percent level and remarkably similar in magnitude, around $1,200.

Table III.2. Annual impacts on earnings: Data weighted to be nationally representative

			Administrative tax data	
	Survey	NDNH	W-2	W-2 and 1099- MISC
Impacts of WIA-funded training				
Earnings: ($)				
Year 1	-1,066	-918	-1,447	-1,473
	(0.381)	(0.222)	(0.112)	(0.106)
Year 2	n.a.	-4	-505	-572
		(0.995)	(0.367)	(0.331)
Impacts of WIA-funded intensive services				
Earnings: ($)				
Year 1	2,710*	1,558	2,418	2,544*
	(0.029)	(0.079)	(0.054)	(0.046)
Year 2	n.a.	1,271*	1,787*	1,719*
		(0.048)	(0.022)	(0.016)
Impacts of WIA-funded training and intensive services				
Earnings: ($)				
Year 1	1,644*	640	971	1,070
	(0.023)	(0.312)	(0.072)	(0.057)
Year 2	n.a.	1,267	1,282	1,147
		(0.080)	(0.057)	(0.052)
Sample size	**4,699**	**33,773**	**34,250**	**34,250**

Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys, Manoli and Patel (2018), and National Directory of New Hires.

Note: Dollars are 2012 dollars. Estimates include zeroes for those who were not employed in the corresponding time period. Estimated impacts are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. Reported *p*-values for impacts are in parentheses and are based on two-tailed t-tests.

n.a. = Not available.

*Significantly different from zero at the .05 level.

There were no statistically significant impacts of WIA-funded training on earnings in the first or second year after random assignment. The estimates were similar in magnitude across the survey, NDNH, and administrative tax data (both W-2 alone and W-2 plus 1099-MISC).[16]

[16] Using the unweighted data, impacts were generally not statistically significant, aside from a negative impact of WIA-funded training on earnings in Year 3, according to administrative tax data (Appendix Table A.5).

D. Comparing average earnings and impacts on earnings using quarterly data from survey and NDNH

Survey-based quarterly earnings estimates exceeded NDNH estimates throughout most of the follow-up period, with differences especially pronounced in the later quarters (Figure III.4). The data sources roughly agreed on customers' earnings in the first calendar quarter after random assignment. Average earnings in Quarter 1 were slightly larger for the NDNH ($1,906) than the survey ($1,880). However, by Quarter 9, survey-reported earnings had climbed to $5,166, whereas NDNH-reported earnings had increased to $4,527.[17] The differences in the estimates of average earnings exceeded 10 percent of reported earnings and were statistically significant for the quarters later in the study period (Quarters 5 to 9) but were not generally statistically significant during the earlier quarters.

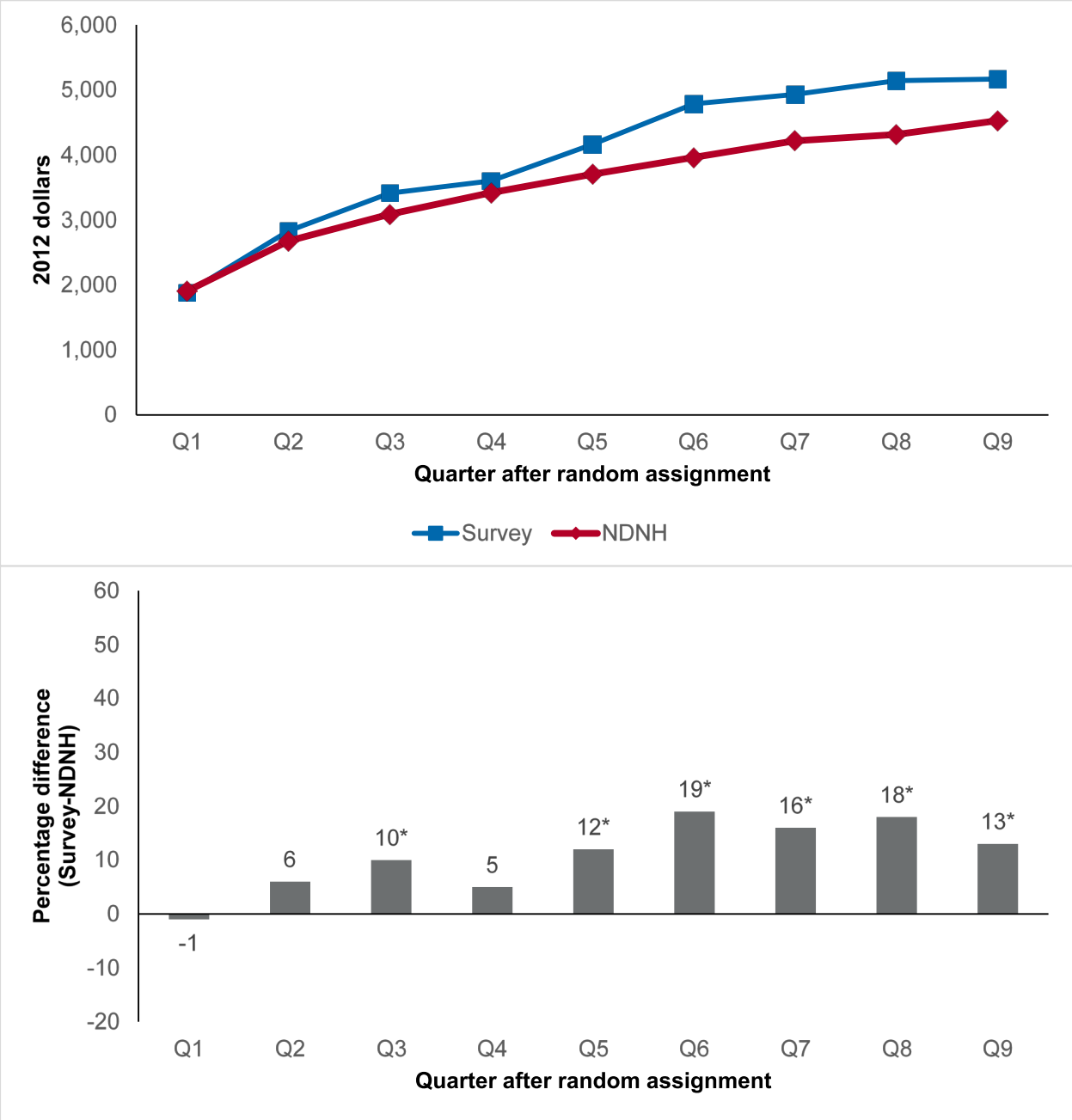
The survey and NDNH data produced similar estimates of the quarterly impact of providing WIA-funded training (Figure III.5). Neither data source indicated that WIA-funded training significantly increased earnings in any quarter. And, in all quarters except for Quarter 9, the impact estimates produced by the different data sources were not statically significantly different from each other (bottom panel of figure).

Estimates of impacts of WIA-funded intensive services were larger when using the survey data than the NDNH data (Figure III.6). Nevertheless, the two data sources yielded a similar qualitative conclusion: WIA-funded intensive services increased earnings in the second year after random assignment (consistent with the annual estimates in Table III.2). Using survey data, we estimated that access to WIA-funded intensive services increased earnings from Quarters 4 to 9 after random assignment, with significant impacts ranging from $623 to $971. In the NDNH data, impacts were smaller and were significant only in Quarters 5 (impact of $526) and 8 (impact of $310). Moreover, the impacts estimated by the two data sources significantly differed from one another in five of nine quarters (bottom panel of figure).

Estimates of impacts of providing WIA-funded intensive and training services were larger when using the survey data than the NDNH data (Figure III.7). As with intensive services alone, the two sources yielded similar qualitative conclusions but differed in their estimated magnitude. Using survey data, we estimated that access to WIA-funded intensive and training services together increased earnings from Quarters 5 to 8 after random assignment, with significant impacts ranging from $550 to $681. The impact in Quarter 9 estimated using survey data was also large ($696) but just missed the threshold for statistical significance. In the NDNH data, impacts were smaller and statistically significant only in Quarter 5 (impact of $543). Moreover, the impacts estimated by the two data sources differed significantly from each other in Quarters 6 to 9 (bottom panel of figure).

[17] Because of rounding to the nearest dollar, arithmetic may appear to be off. More precise values are presented in the appendix tables.

Figure III.4. Earnings estimated using NDNH versus survey data

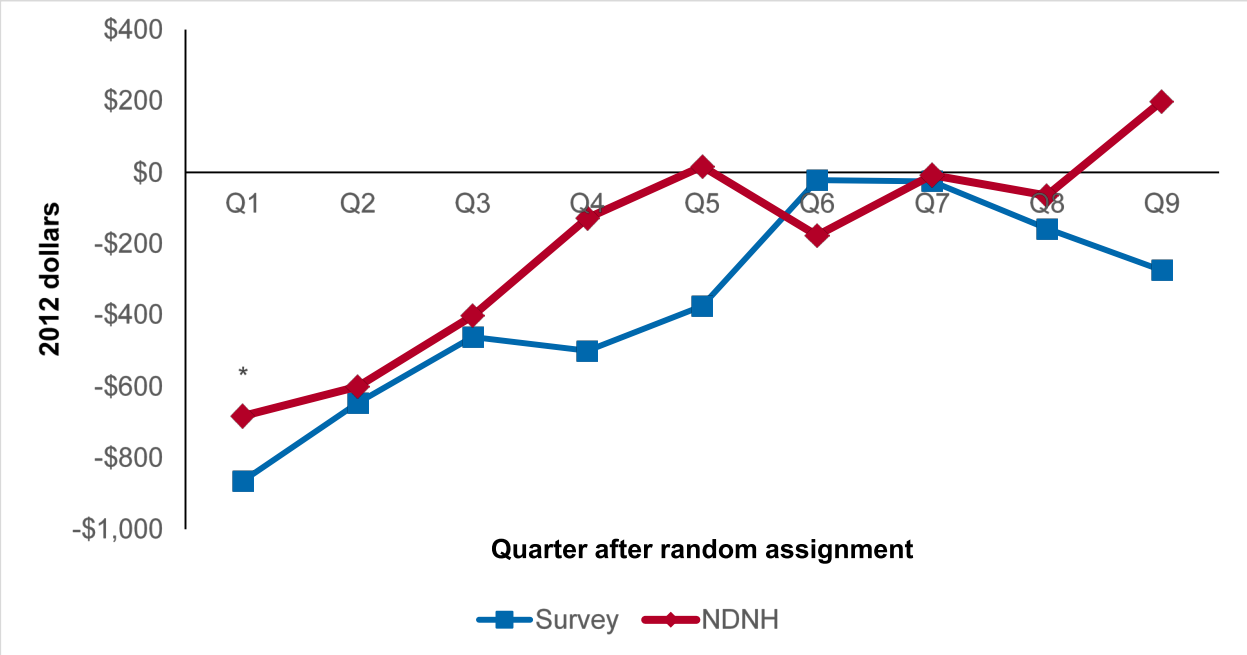


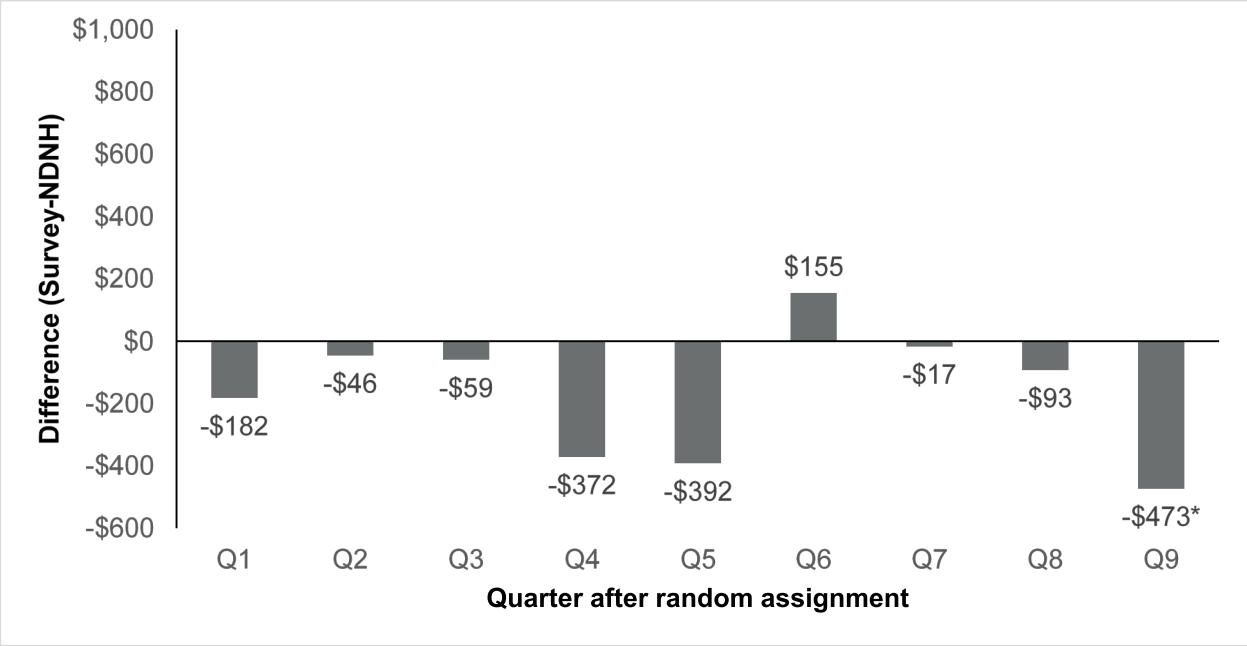
Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimates include zeroes for those who were not employed in the corresponding time period. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.6.

*Significantly different from zero at the .05 level.

Figure III.5. Impact of WIA-funded training



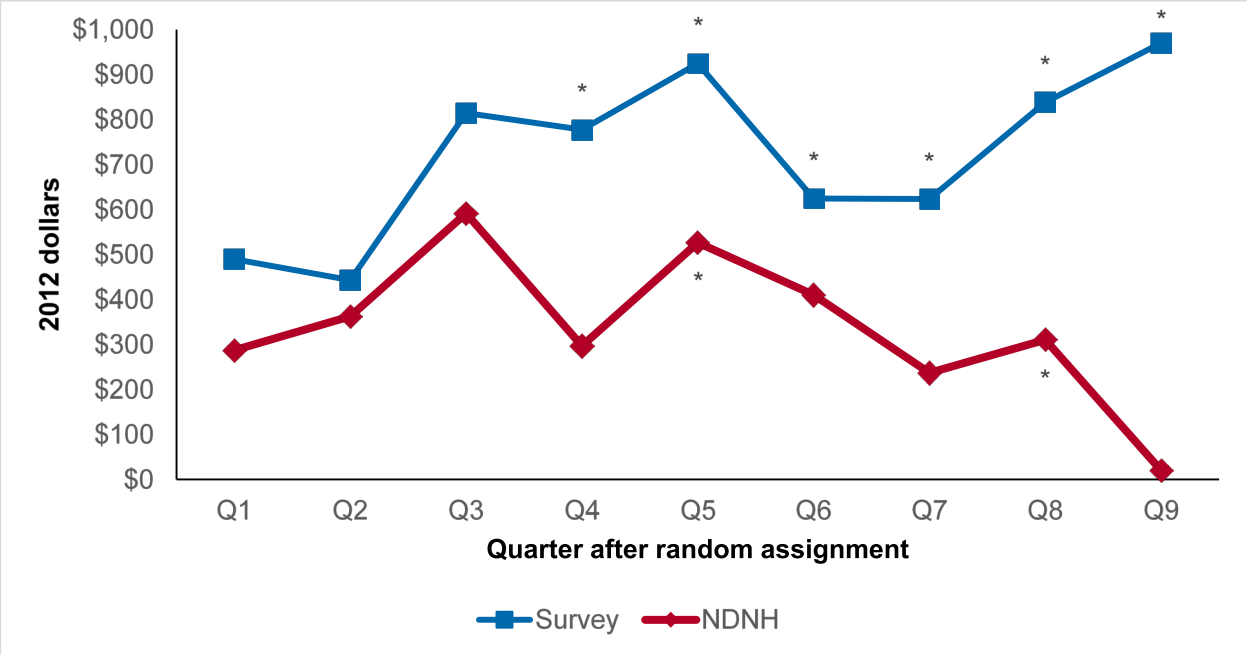


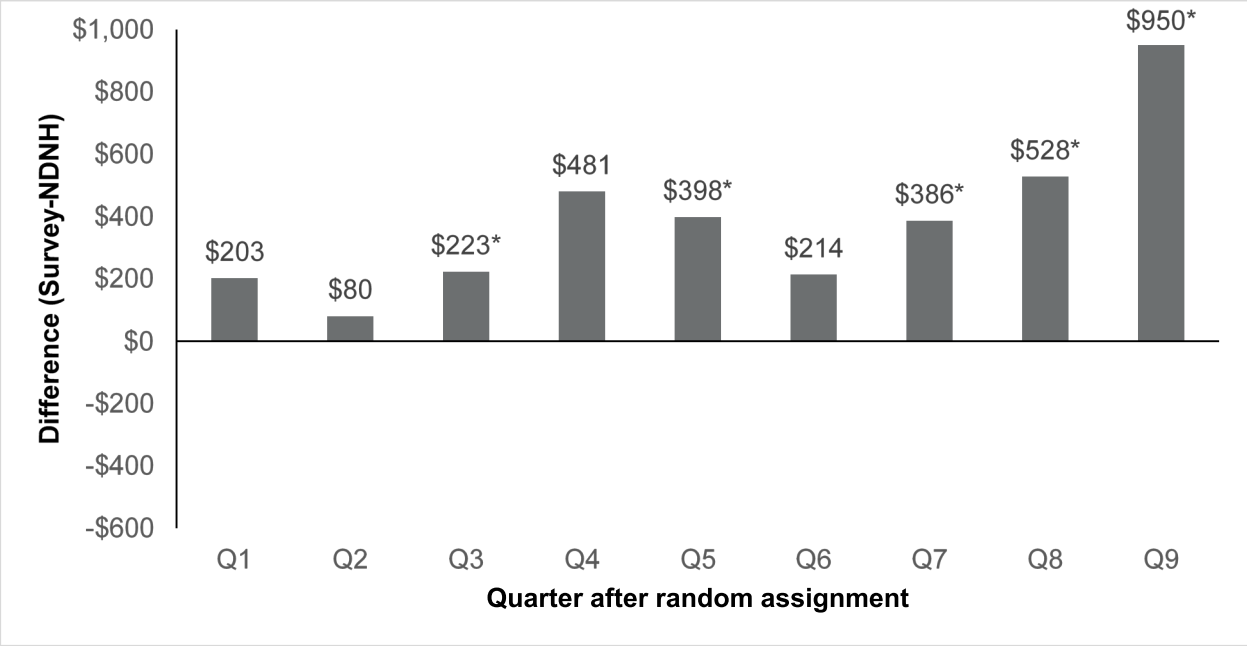
Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimates include zeroes for those who were not employed in the corresponding time period. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey and (6) that the job-seeker completed the survey. For further details, see Appendix Table A.7.

*Significantly different from zero at the .05 level.

Figure III.6. Impact of WIA-funded intensive services



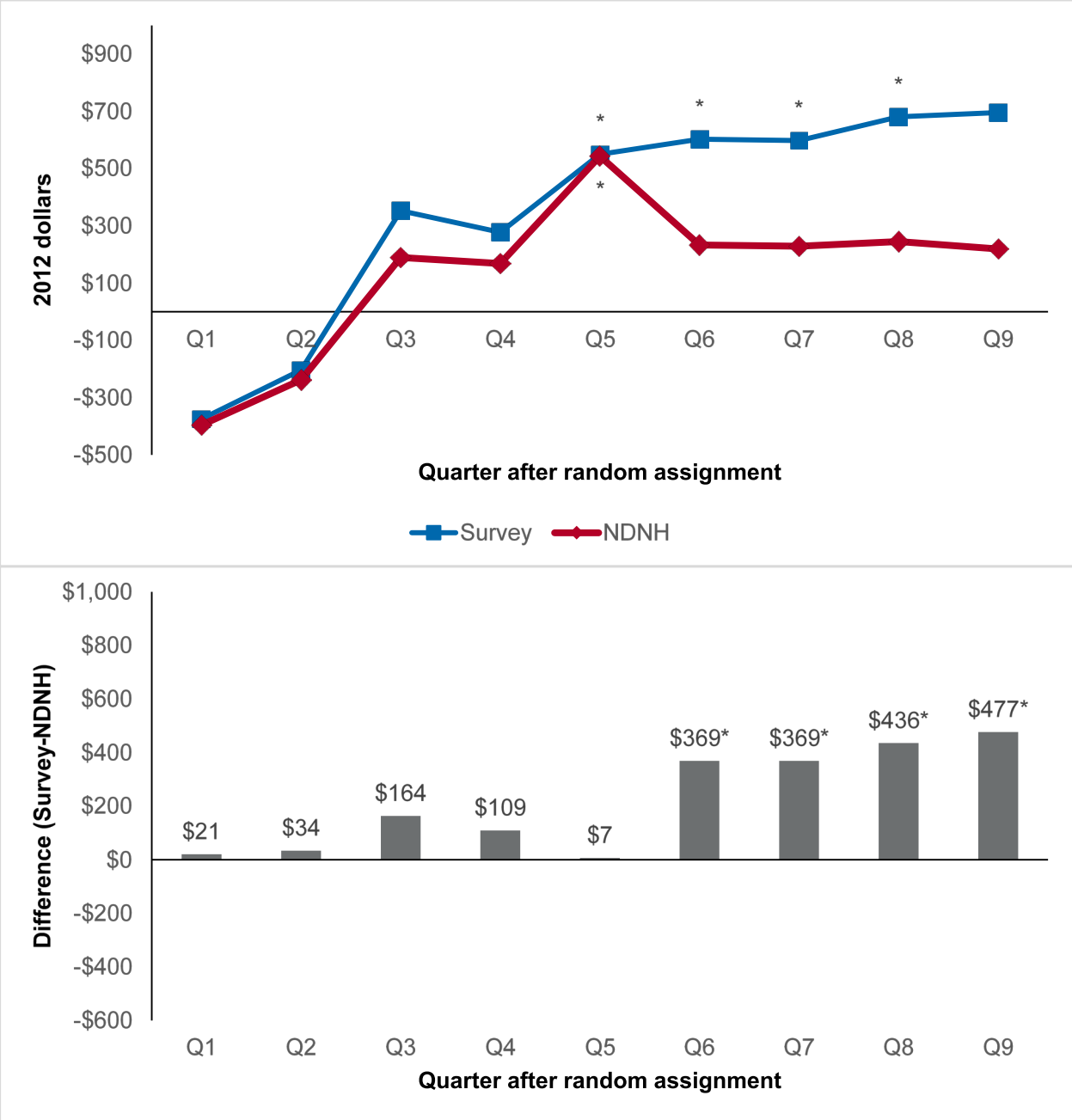


Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimates include zeroes for those who were not employed in the corresponding time period. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey and (6) that the job-seeker completed the survey. For further details, see Appendix Table A.7.

*Significantly different from zero at the .05 level.

Figure III.7. Impact of WIA-funded intensive and training services



Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimates include zeroes for those who were not employed in the corresponding time period. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey and (6) that the job-seeker completed the survey. For further details, see Appendix Table A.7.

*Significantly different from zero at the .05 level.

IV. INVESTIGATING DIFFERENCES IN QUARTERLY DATA FROM THE SURVEY AND NDNH

The analyses presented in Chapter III indicate important differences in the impacts estimated across the data sources, particularly when using the finer-grained quarterly data available in both the surveys and NDNH. In this chapter, we delve deeper into the reasons for the differences. We focused our analysis on understanding differences in mean quarterly earnings across the two data sources and use the differences in means to assess the sources' relative strengths and drawbacksin estimating the impact of the WIA Adult and Dislocated Worker programs.[18] Section A describes our approach for this analysis and Section B contains the results.

A. Approach to investigating differences across the survey and NDNH

Although we applied the same methods to estimate impacts using the survey, NDNH, and administrative tax data, the differences in outcome measurement, timing, coverage, and other factors described in Section II resulted in different point estimates and statistical significance levels. We used a simple framework to understand the reasons for the different estimates and develop our analytic approach for investigating them (Figure IV.1). Differences in earnings estimates can be decomposed into differences in samples and differences in the underlying data. Holding samples constant, differences in earnings can be further decomposed into differences in the number of jobs reported and the earnings reported for each job. Finally, we can draw an additional distinction between differences in the number of jobs caused by the limited coverage of the NDNH and the limited accuracy of the survey data (Chapter II, Section E).

We took a three-step approach to conduct an in-depth investigation of the differences in estimates of average quarterly earnings derived from the survey and NDNH data, corresponding to the three-tiered nature of the factors identified in Figure IV.1. Within each tier, we considered which factors substantively contributed to differences in mean earnings estimates.

We focused on probing the differences underlying mean earnings and employment rates—rather than the impacts themselves—because these differences can interact in different ways to affect the impact estimates. For instance, a larger earnings impact on the survey compared with NDNH could reflect overreporting of earnings on the survey, or less coverage of informal labor in the NDNH, or a combination of the two.

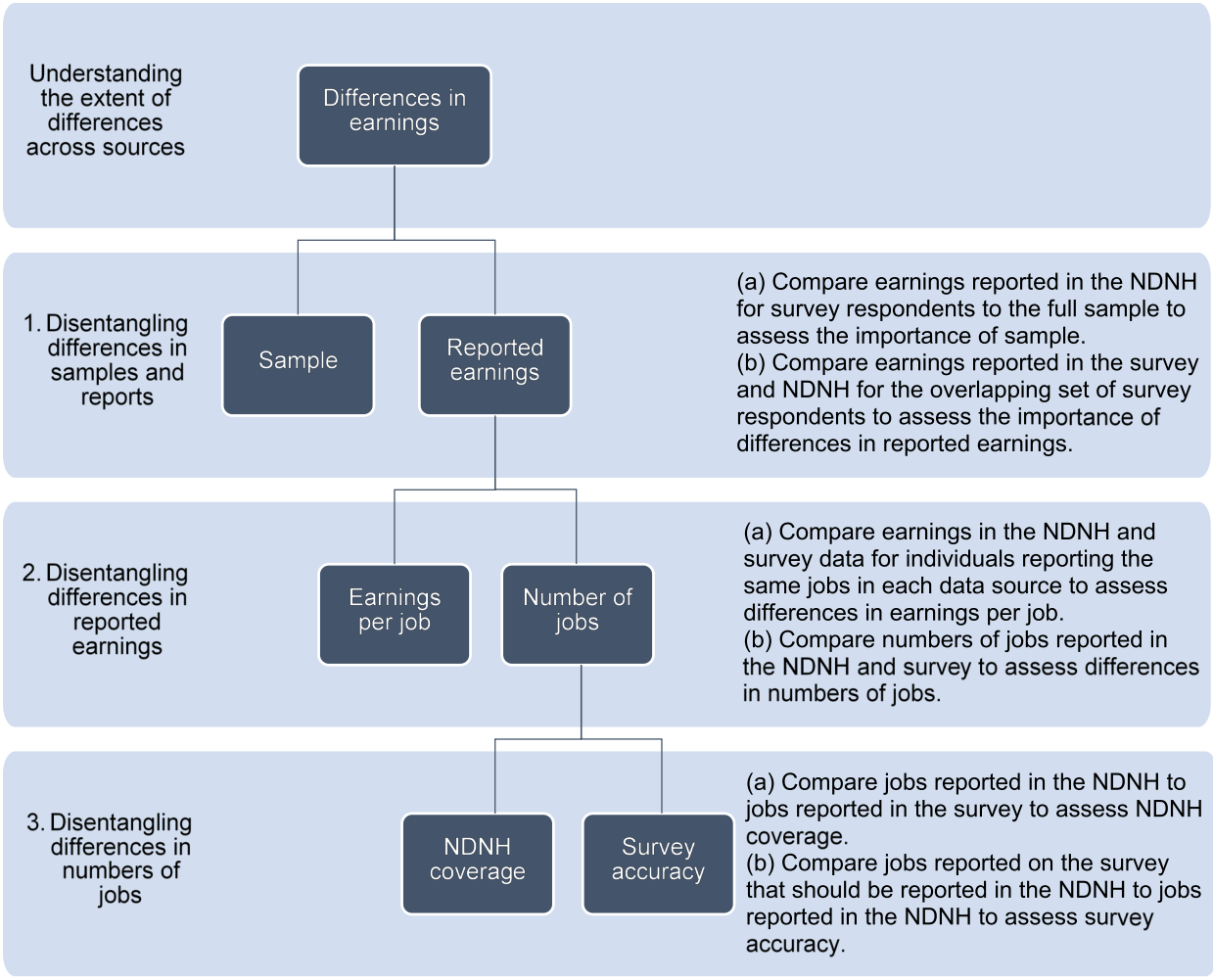
1. Disentangling differences in samples and reports

If there were no individual-level reporting gaps (for example, some types of labor not covered in certain data sources) or errors (for example, recall on the survey, underreporting to sources of administrative data), different estimates of average earnings across the two data sources could result if the sample covered by each differed; subsampling for the survey and survey nonresponse could be reasons the samples would differ. Conversely, if the samples in the

[18] This more in-depth analysis could not be done for comparisons including the administrative tax data because the number of jobs held by an individual—which is crucial to our approach—could not be estimated using that data source. In addition, we could not link the administrative tax data directly to the NDNH, due to restrictions on the use of both data sets. Finally, administrative tax data contain only annual, not quarterly, data.

two data sources were identical, different estimates of average earnings could result from reporting gaps or errors.

Figure IV.1. Understanding differences in earnings across data sources



To disentangle these explanations, we first compared estimates based on NDNH data for all study participants with those based on NDNH data only for survey respondents (Step 1.a in Figure IV.1). Differences in these estimates would stem from the differences in analytic samples and, in particular, to the possibility of survey nonresponse for a nonrandom subset of individuals.

Next, we compared estimates based on survey data to those from NDNH data only for survey respondents (Step 1.b in Figure IV.1). Holding the samples constant across the two enabled us to isolate differences attributable to reporting differences or errors.

For these analyses, we used *t*-tests to determine whether each of the examined differences between the pairs of estimates were statistically significant. Because of the number of comparisons we made, there is a large probability that some comparisons would meet conventional thresholds for statistical significance by chance even if no true differences underlay the estimates. However, given the exploratory nature of this analysis, we did not want to miss

interesting trends that would merit further study. Therefore, we did not adjust our statistical significance tests for the multiple comparisons we conducted.

All estimates further accounted for the details of the study's design (Chapter II, Section A and Rotz et al. 2017), as well as the correlations between the different values compared.[19] Throughout this analysis, we had to make one additional assumption. To protect the anonymity of NDNH records, we could only link individuals' survey-reported earnings ventile—not exact survey-reported earnings–with NDNH data. Ventiles are the 20 groups that result when we divide the distribution of earnings into 20 parts, with each containing one twentieth of the study sample. As a result, many of our comparisons required us to assume that the correlation between NDNH-reported earnings and survey-reported earnings was the same as that between NDNH-reported earnings and the earnings ventiles from the survey.

2. Disentangling differences in reported earnings

Unlike differences in estimated earnings caused by differences in samples, disagreements between data sources in the earnings reported for a fixed sample do not clearly fault either source. The differences only show there is a limitation with at least one of the data sources and provide no indication of the relative reliability of either one. We therefore further disaggregated differences between reported earnings across the two data sources into two components: disagreements in the earnings associated with a given job and those in the number of jobs held.

To understand how reported earnings in a given job differed across the survey and NDNH data, we estimated average NDNH-reported earnings and survey-reported earnings *for the same job* in Quarters 1 and 9 after random assignment (Step 2.a in Figure IV.1). The NDNH does not contain enough information for us to match specific jobs across the data sources. Therefore, we restricted our analysis to survey respondents who on the survey reported holding exactly one job and who in the NDNH data also appeared to have held exactly one job. However, because the survey responses also include types of jobs that are not likely to be reportable in the UI wage records (and, hence, would not be expected to be found in the NDNH), we further restricted the analysis to survey respondents whose single job would be expected to be reported to UI wage records, which we refer to as NDNH-reportable jobs.

Because UI laws and definitions of job characteristics are nuanced, we could not identify NDNH-reportable jobs on the survey with certainty. Therefore, we used four plausible working definitions of an NDNH-reportable job, as described in Table IV.1. We focused our analyses on the simplest and most plausible of these: our baseline definition considers a job reportable to the NDNH if the survey respondent identified it as a regular job (rather than a temporary position, on-call work, day labor, contracting, or self-employment).[20]

[19] When comparing the NDNH and survey data, we also omitted (1) observations for which NDNH data were not available and (2) observations for individuals who responded to the survey but did not have observed data on earnings. This small reduction in sample size reduced the potential explanations for differences across data sources.

[20] We also conducted this analysis using a slightly larger sample of individuals who held no non-reportable jobs in the survey, two or more reportable jobs in the survey, and the same number of jobs in the NDNH and survey data. This produced similar results.

Table IV.1. Definitions of NDNH-reportable jobs based on survey-reported characteristics

Definition	Criteria and comparison to the baseline definition
Definition 1: Baseline	Includes all regular part- or full-time jobs (that is, jobs not classified as self-employment, independent contracting, contracting,[a] temporary jobs, day labor, and on-call labor).
Definition 2: Least inclusive	Includes all regular part- or full-time jobs (that is, jobs not classified as self-employment, independent contracting, contracting,[a] temporary jobs, day labor, and on-call labor) **offering at least one of the following benefits: health insurance, paid vacations, paid sick days, paid holidays, or pensions.** We added the restriction to only jobs offering benefits to focus on cases in which the jobs reported by respondents do not reflect permanent employment.
Definition 3: More inclusive	Includes all regular part- or full-time jobs (that is, jobs not classified as self-employment, independent contracting, contracting,[a] temporary jobs, day labor, and on-call labor), **as well as contractor and temporary jobs offering at least one of the following benefits: health insurance, paid vacations, paid sick days, paid holidays, or pensions**. We added this addition of contracting and temporary positions because such positions might be reported to the NDNH, but limited to those that offer benefits because these are the most likely to be reported.
Definition 4: Most inclusive	Includes all regular part- or full-time jobs (that is, jobs not classified as self-employment, independent contracting, contracting,[a] temporary jobs, day labor, and on-call labor), **as well as contractor, temporary, day labor, and on-call jobs offering at least one of the following benefits: health insurance, paid vacations, paid sick days, paid holidays, or pensions**. We added contracting, temporary jobs, day labor, and on-call labor because such positions might be reported to the NDNH, but limited to those that offer benefits because these are the most likely to be reported.

[a] Contractor positions, or jobs with a contracting agency, could be included in the NDNH. The survey distinguished these jobs from independent contracting jobs, which are excluded.

To further understand the potential causes for differences in reported earnings in the same job, we organized the customers into three groups: one with NDNH-reported earnings less than 90 percent of survey-reported earnings, another with NDNH-reported earnings greater than 110 percent of survey-reporting earnings, and a third group in which the two sources roughly agreed. Across these groups, we compared the average hourly wage rate, average hours worked per week, likelihood of working in each of the top-five broad industry categories, and likelihood of working in each of the top-five occupations.

An important caveat applies to our analysis of earnings per job. Because we must limit the analysis to customers with at least one job and with no jobs that should not be reported to the NDNH, our sample is not representative of all employed customers, let alone the full sample of customers studied. It could be that the kinds of jobs not captured by the NDNH are subject to different recall and reporting patterns than are regular jobs.

We also compared the total number of jobs reported in the NDNH and the survey for survey responders (Step 2.b in Figure IV.3) to understand the importance of differences in how many jobs each data source captured.

3. Disentangling differences in numbers of jobs

Two factors could drive differences in the number of reported jobs: (1) survey respondents holding jobs such as farm labor or self-employment not included in the NDNH (or not reported to UI by employers) or (2) customers forgetting or intentionally omitting jobs in their survey

responses. Factor (2) is likely a more serious issue for jobs held earlier in the study period than for those later in the period, whereas Factor (1) should remain the same throughout the follow-up period. Therefore, differences between the NDNH and survey data over time (for Quarters 1 and 9) can help us understand the relative magnitudes of Factors (1) and (2) as explanations for the differences in numbers of jobs.

To better understand Factor (1), we compared the average number of total jobs reported in the survey to the average number of NDNH-reportable jobs reported in the survey (Step 3.a in Figure IV.1). As in our analyses of earnings holding the number of jobs constant, we focused on the baseline definition for NDNH-reportable jobs and used other definitions to confirm that the definition we preferred did not unduly affect our findings.

To determine the extent to which Factor (2) was an issue, we compared the average number of jobs reported in the NDNH to the average number of NDNH-reportable jobs reported in the survey (Step 3.b in Figure IV.1). We further examined how the relationship between these two measures of number of jobs changed over time, comparing differences in the number of jobs reported for Quarters 1 and 9 after random assignment.

In contrast to understanding differences in earnings reports for a specific job, estimating differences in the number of jobs held between the NDNH and survey data was straightforward. We simply compared the averages in each data source. We conducted this analysis for Quarters 1 and 9 to determine if differences varied over the study period. We used the results to estimate the proportion of reportable jobs survey respondents forgot and the proportion of jobs likely to be informal in nature. We compared these for Quarters 1 and 9 to determine whether survey recall was indeed a larger contributor to the differences early in the follow-up period (that is, longer in the recall period).

B. Causes for differences in quarterly data

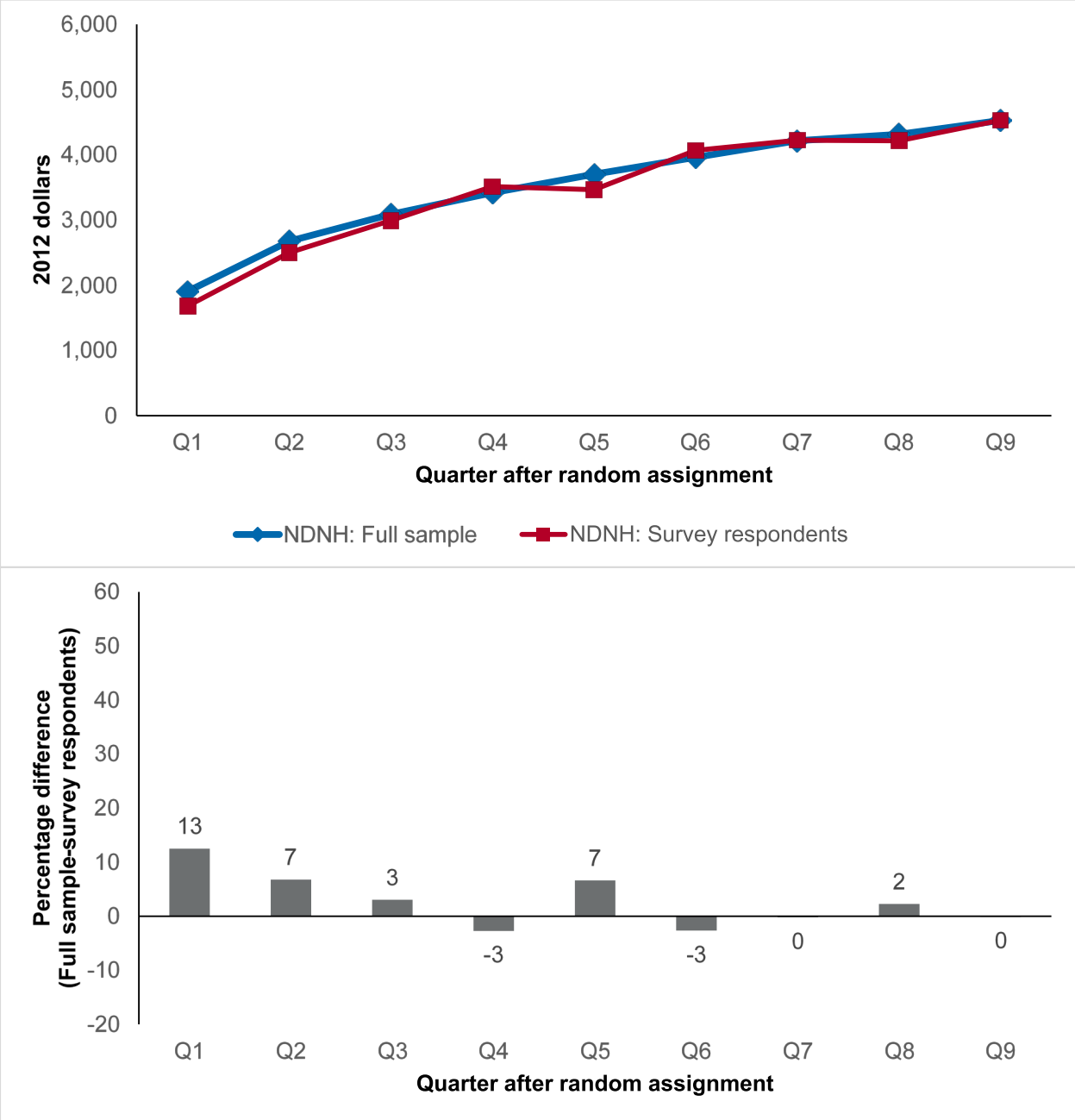
We found that three factors likely explained most of the differences between the two data sources: (1) many survey respondents worked in jobs not captured by the NDNH, such as informal, day labor, self-employment, and independent contracting; (2) survey respondents typically overreported their earnings in any given job, especially early in the follow-up period; and (3) survey respondents typically underreported the number of jobs they held early in the follow-up period. This section presents our findings, following the outline presented in Figure IV.1.

1. Disentangling differences in samples and reports

a. Differences in samples

Differences in earnings estimated using survey and NDNH data appeared to stem from differences in the data collected and not differences in the samples covered (Figure IV.2). Comparing estimates based on the full NDNH sample with estimates calculated using NDNH data for only survey respondents suggested that randomly selecting the survey sample and using nonresponse weights successfully eliminated any meaningful differences in reported earnings between the full study sample and the sample of survey respondents. The conditional differences between the full-sample estimates and the survey-respondent estimates of average NDNH-reported quarterly earnings were all small, and none were statistically significant. Thus, we

Figure IV.2. Earnings estimated using NDNH data, all study participants versus survey respondents

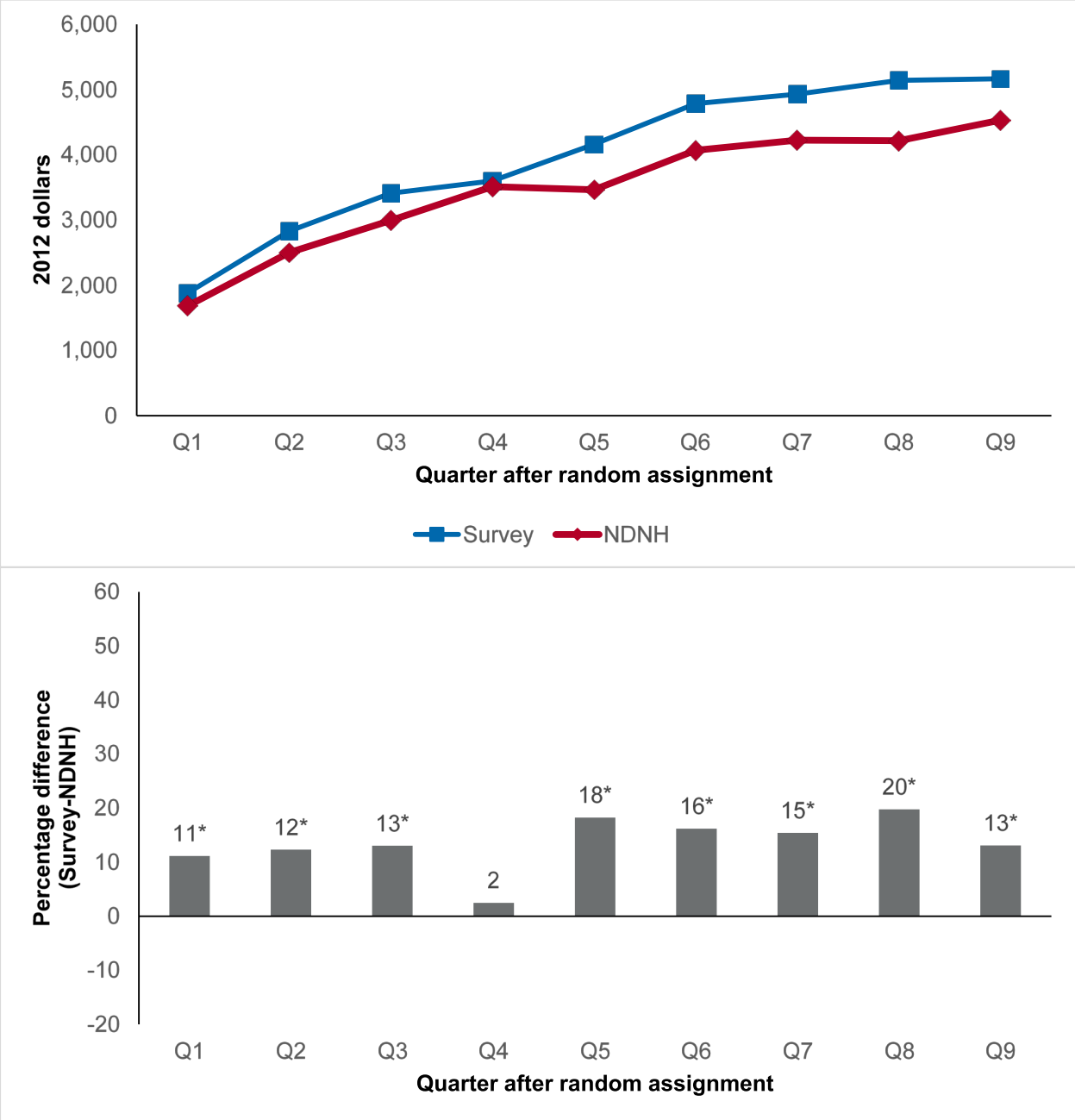


Source: National Directory of New Hires.

Note: Estimates include zeroes for those who were not employed in the corresponding time period. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.6.

*Significantly different from zero at the .05 level.

Figure IV.3. Earnings estimated using sample of survey respondents, NDNH versus survey data



Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimates include zeroes for those who were not employed in the corresponding time period. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.6.

*Significantly different from zero at the .05 level.

concluded that the weights worked appropriately, and we could rule out differences in the samples covered by each data source as an important explanation for differences across the data sources.

We conducted the same analysis for each study group separately, and the results were consistent, with one important exception: late in the study period, in the core-and-intensive group, survey respondents tended to have significantly higher NDNH earnings than nonrespondents (Appendix Table A.6). Survey respondents from the core-and-intensive group reported earnings in Quarters 7 through 9 that were $251 to $341 (or 6 to 7 percent) higher than earnings for all members of this study group. This likely reflects noise in the data, and no evidence suggests the underlying characteristics of the core-and-intensive group, or the services received by this group, drove this trend.[21] However, this finding suggests that differences in the extent that the weights corrected for survey nonresponse across study groups might lead to differences in estimated impacts across data sources.

b. Differences in reports of earnings

Holding constant the sample analyzed, larger differences between the data sources' estimates of average earnings became apparent (Figure IV.3). The data sources consistently disagreed on customers' earnings in the first calendar quarter after random assignment, with data from the survey suggesting higher earnings than the data from the NDNH. The difference ranged from $88 in Quarter 4 to $926 in Quarter 8. Moreover, in every quarter except for the fourth after random assignment, the difference was $198 or more and statistically significant.

2. Disentangling differences in reported earnings

a. Earnings in a given job

Survey-reported earnings in a given job exceeded those in the NDNH for the same job early in the follow-up period (Table IV.2). In Quarter 1, customers with exactly one reportable job in both data sources (presumably the same job) reported earnings that were $661 higher in the survey than in the NDNH, a statistically significant difference of 16 percent. This pattern held under alternative definitions of reportable jobs and when expanding the sample to include customers with more than one reportable job (but no nonreportable jobs and the same number of jobs across data sources). This suggests that individuals overreported job-specific earnings in survey data when asked to recall information about jobs held early in the study period (Table IV.2).[22]

[21] As described in Rotz et al. (2017), the WIA Gold Standard Evaluation study team developed nonresponse weights by modeling survey response based on demographic characteristics within study groups. Therefore, the weights accounted for the potential for different demographic characteristics to predict survey response across study groups. Moreover, the different study groups were equivalent at baseline and survey response rates did not vary significantly across study groups. Given this, and the rigorous random assignment study design, it seems highly likely that the differences simply reflect statistical noise.

[22] Note that although the way in which the WIA Gold Standard Survey asked customers to report on their wage rate might have led to some overreporting of earnings early in the follow-up period, it is unlikely that this drove the entire observed difference between the data sources. In particular, the survey asked respondents to report the wage they currently received at their jobs (or the last wage received for jobs that had ended). This could lead to

Table IV.2. Earnings in reportable jobs, individuals with a single, reportable job in the survey and a single job in the NDNH

	Definition of a job reportable in the NDNH			
	Definition 1[a]	Definition 2[b]	Definition 3[c]	Definition 4[d]
Quarter 1 earnings: ($)				
Survey[e]	4,590	4,024	4,721	4,487
NDNH	3,930	3,983	4,023	3,893
Difference	661*	678*	697*	594*
	(0.000)	(0.000)	(0.000)	(0.000)
Sample size	**1,030**	**1,011**	**1,159**	**1,203**
Quarter 9 earnings: ($)				
Survey[e]	6,916	7,325	6,835	6,765
NDNH	6,695	7,249	6,652	6,588
Difference	221	76	183	177
	(0.155)	(0.655)	(0.203)	(0.217)
Sample size	**2,073**	**1,850**	**2,165**	**2,200**

Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Dollars are 2012 dollars. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.8.

[a] Jobs respondents categorized on their survey as regular jobs, excluding those described elsewhere on the survey as self-employment.

[b] Jobs respondents categorized on their survey as regular jobs that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions. Excludes all jobs described as self-employment.

[c] Jobs respondents categorized on their survey as contractor or temporary positions that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions; or as regular jobs. Excludes all jobs described as self-employment. Contractor positions, or jobs with a contracting agency which the NDNH might include, were distinguished on the survey from independent contracting, which is excluded.

[d] Jobs respondents categorized on their survey as contractor, temporary, day labor, or on-call positions that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions; or as regular jobs. Excludes all jobs described as self-employment. Contractor positions, or jobs with a contracting agency which the NDNH might include, were distinguished on the survey from independent contracting, which is excluded.

[e] Estimated using the mean survey-reported earnings of customers in the same ventile of survey-reported earnings.

*Significantly different from zero at the .05 level.

Survey-reported earnings in a given job were roughly in line with those in the NDNH for the same job late in the follow-up period (Table IV.2). Using the preferred definition of a job reportable in the NDNH, customers with exactly one reportable job in both data sources had similar average Quarter 9 earnings in the survey ($6,916) and the NDNH ($6,695). The

overstating earnings early in the follow-up period if customers received raises throughout their time at a job. However, customers in the study sample did not often hold jobs for several years at a time, suggesting any raises received would be small. Among those who worked in the first quarter after random assignment, only about 57 percent reported holding the same job in the fifth quarter after random assignment and only about 35 percent reported holding the same job at the time of the 30-month follow-up survey. This suggests limited scope for this type of wage increase.

difference between the two estimates was not statistically significant. This held under alternative definitions of reportable jobs and when expanding the sample to include individuals with more than one reportable job.

We conducted the same analysis for each study group separately and the results were consistent, with one important exception: those in the core group who held exactly one reportable job according to both the survey and NDNH data reported average Quarter 9 earnings that were significantly higher in the NDNH than in the survey data (Appendix Table A.8). This could have led to differences across data sets in the impacts estimated by the WIA Gold Standard Evaluation.

The difference between the data sources in reported earnings for a given job early in the reporting period, but lack thereof in later periods, suggests recall error drove the early differences. We would expect recall error to be a larger issue when individuals report about earnings in jobs held further in the past. In contrast, an alternative explanation for survey-reported earnings exceeding NDNH-reported earnings (systematic underreporting of earnings to UI agencies to avoid UI contributions) would not change dramatically over the period studied. Furthermore, for the WIA Gold Standard Evaluation at least, individuals' overreporting of job-specific earnings in survey data seems to have been attributable to recall error early in the follow-up period, and not the way in which earnings measures were constructed (which was constant over the follow-up period).

Specific job characteristics do not appear to explain differences in earnings in a given job in Quarter 9 (Table IV.3). The only characteristic that seemed consistently associated with the difference between NDNH- and survey-reported earnings per job was working in the manufacturing industry (Table IV.3). Only 3 percent of individuals with lower survey than NDNH earnings, but 9 percent of individuals with similar earnings in both data sets and 22 percent of individuals with higher NDNH than survey earnings for the same job worked in manufacturing. The differences between each of the proportions were statistically significant. This industry is typically associated with more formal employment and substantial amounts of overtime work (Bureau of Labor Statistics 2017). The pattern therefore suggests that the WIA Gold Standard Evaluation survey might underestimate earnings for jobs when overtime work accounts for a large proportion of earnings. Indeed, the survey asked respondents to provide a single measure of hourly wages and a single measure of hours worked per week for each job they reported on the survey.

The data on average hourly wages provided some evidence that either workers with higher wage rates tended to underreport their earnings on the survey or workers with lower wage rates tended to overreport their earnings (Table IV.3). Customers with greater NDNH-reported earnings than survey-reported earnings indicated on the survey that they had higher hourly wages ($15.11) than those who had comparable earnings across the data sources ($13.86). This difference was statistically significant. Customers with greater survey-reported earnings than NDNH-reported earnings had hourly wages of $13.52, which was statistically indistinguishable from those who had comparable earnings across the data sources.

Table IV.3. Quarter 9 job characteristics among customers with exactly one job in the NDNH and one reportable job in the survey

	Means			Conditional differences	
	Survey earnings exceed NDNH (1)	Earnings match (2)	NDNH earnings exceed survey (3)	1 – 2	2 – 3
Average hourly wages ($)	13.52	13.86	15.11	-0.34 (0.664)	-1.26* (0.014)
Average hours per week	42.01	40.57	42.01	1.44 (0.064)	1.68 (0.086)
Most frequent industries: (%)					
Health care and social assistance	0.271	0.220	0.210	0.051 (0.337)	0.011 (0.867)
Manufacturing	0.028	0.090	0.216	-0.062* (0.004)	-0.126* (0.021)
Accommodation and food services	0.057	0.101	0.061	-0.044 (0.254)	0.040 (0.256)
Retail	0.075	0.059	0.072	0.016 (0.474)	-0.013 (0.633)
Administrative and support, waste management, and remediation services	0.051	0.106	0.039	-0.054 (0.084)	0.066* (0.030)
Most frequent occupations: (%)					
Nursing, psychiatric, and home health aides	0.092	0.041	0.076	0.051* (0.025)	-0.035 (0.133)
Retail sales workers	0.089	0.082	0.027	0.007 (0.871)	0.055 (0.132)
Information and record clerks	0.068	0.053	0.055	0.015 (0.369)	-0.002 (0.906)
Material-moving workers	0.062	0.028	0.055	0.034 (0.249)	-0.027 (0.234)
Motor vehicle operators	0.050	0.053	0.036	-0.003 (0.853)	0.018 (0.322)
Sample size	**707**	**688**	**678**		

Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Dollars are 2012 dollars. Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.9.

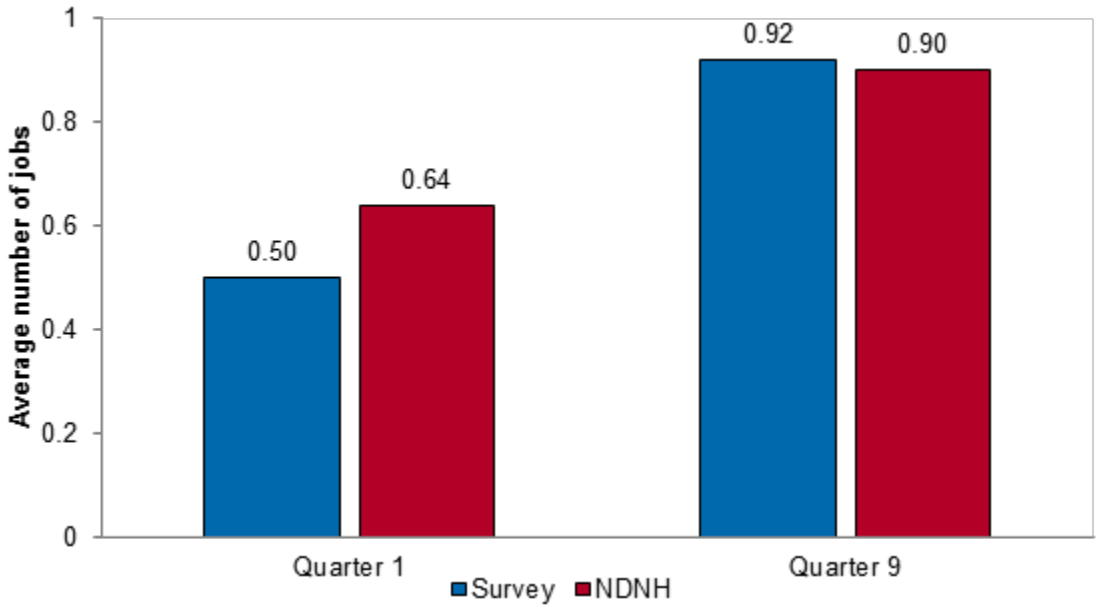
*Significantly different from zero at the .05 level.

b. Number of jobs

Customers reported fewer jobs on the survey than were recorded in the NDNH early in the study period. But this reversed over time (Figure IV.4). The number of jobs customers

reported holding in the survey was lower than that calculated from the NDNH in Quarter 1, but this difference had eroded by Quarter 9. According to NDNH data, the average customer held 0.64 jobs in the first calendar quarter after random assignment, compared with 0.50 jobs reported for the quarter in the survey data. By the ninth quarter, the relationship had reversed and survey data indicated that customers held 0.92 jobs on average, against the 0.90 captured in the NDNH. This suggests underreporting of jobs in surveys could be particularly problematic for longer recall periods.

Figure IV.4. Number of jobs reported for survey respondents, survey versus NDNH data



Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.10.

Because we estimated employment based on whether an individual reported any jobs held, it is unsurprising that we found a similar pattern in overall employment rates. Employment rates were initially lower in the survey data than in the NDNH but this pattern quickly reversed (Figure IV.6). According to survey data, 44 percent of customers were employed in the first quarter after random assignment. By comparison, NDNH data for survey respondents implied an employment rate of 50 percent during this period. By the ninth quarter, the survey-reported employment rate was 79 percent, whereas the employment rate implied by the NDNH data was 72 percent. Differences between the survey estimate and the NDNH estimate were statistically significant in Quarter 1 and, with the opposite sign, from Quarter 3 onward.

Like the results for number of jobs, patterns in the differences between employment rates as estimated by each data source suggest both differences in which jobs were covered and in reporting jobs might have contributed to the differences between survey and NDNH earnings

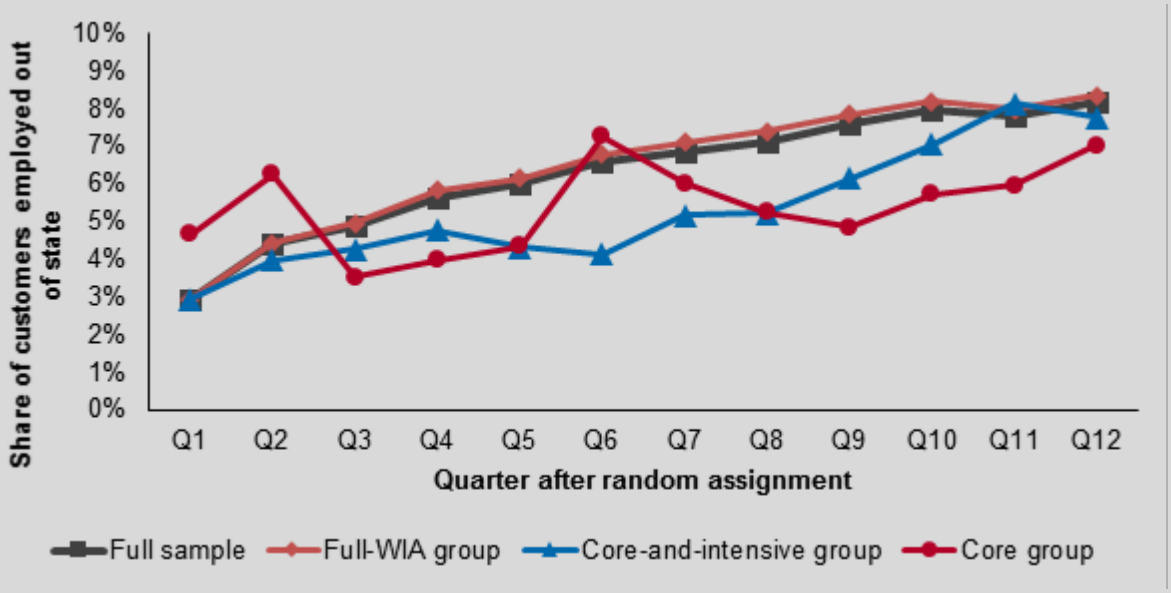
estimates. The difference in employment rates for the later quarters of the follow-up period suggest that the NDNH excluded certain jobs. The opposite-signed difference for the first quarter further suggests the presence of recall error (that is, individuals not reporting certain jobs) in the survey. Finally, examining the trend in the difference over time suggests that recall error was a larger issue than NDNH coverage for the first quarter after random assignment. In the second quarter after random assignment, the two sources of error were roughly balanced. Thereafter, the limited coverage of the NDNH led to lower estimated employment rates.

A note on out-of-state employment

Before the availability of NDNH data, researchers commonly relied on state-specific UI records for administrative data on earnings and employment (Barnow and Greenberg 2015). A common criticism of that approach is that individuals can obtain out-of-state employment not captured by state-specific records, particularly if they live and work in a metropolitan area that straddles state borders. We avoid this concern by using only administrative data from across the United States. Further, our analysis suggests the concern is duly warranted. Over the 12 quarters for which we could obtain NDNH data, rates of out-of-state employment ranged from 3 percent in Quarter 1 to 8 percent in Quarter 12 (Figure IV.5). Moreover, 16 percent of customers in our sample held at least one out-of-state job and the average individual who was employed at any point in the follow-up period was employed out of state in 10 percent of all quarters he or she was employed (not shown).

Although potentially problematic, this would not lead to bias in impact estimates if out-of-state employment was uniform across study groups. But the NDNH indicates that access to WIA-funded training increased the probability of out-of-state employment and the average number of quarters worked out of state. The full-WIA group was generally more likely than either other study group to be employed out of state in a given quarter. In 6 of the 12 quarters examined, the difference between the full-WIA group and the core group was statistically significant. This was true for 2 of the 12 quarters for the comparison between the full-WIA and core-and-intensive groups. This evidence suggests past studies based on UI data that omitted earnings from out-of-state jobs likely underestimated the impacts of WIA-funded training.

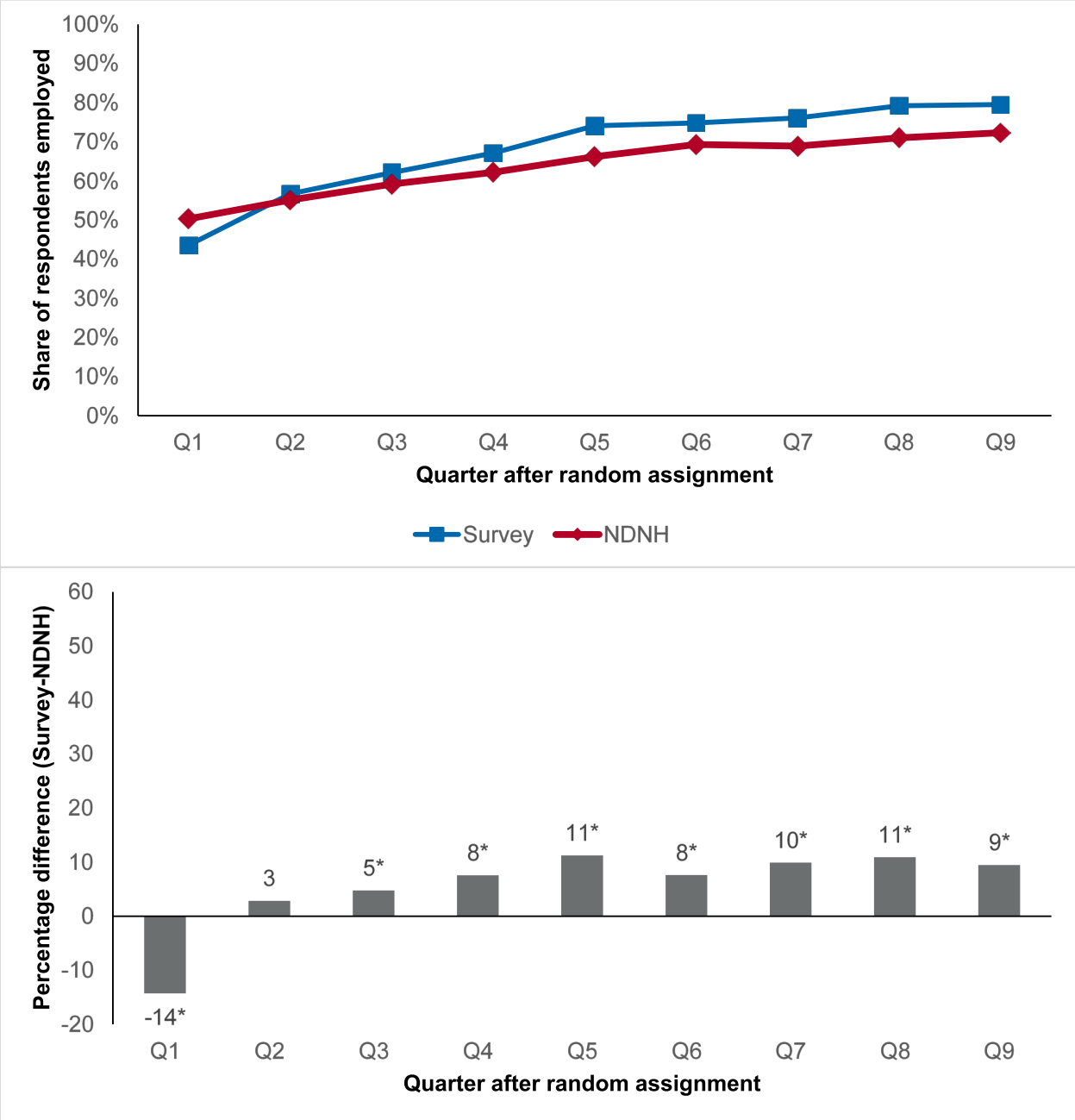
Figure IV.5. Out-of-state employment rates



Source: National Directory of New Hires.

Note: Estimated means are regression-adjusted to account for the administrative region in which customers enrolled in WIA services. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. See Appendix Table A.11 for further details.

Figure IV.6. Employment rates for survey respondents, survey versus NDNH data



Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.12.

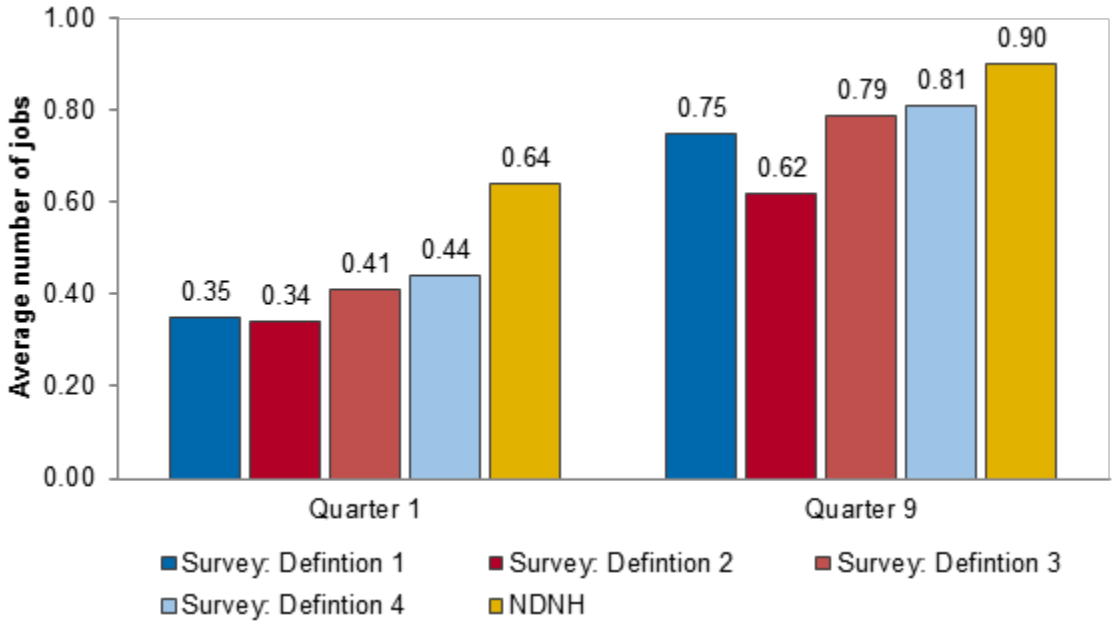
*Significantly different from zero at the .05 level.

3. Disentangling differences in numbers of jobs

a. NDNH coverage

NDNH data were unlikely to capture one-tenth to one-third of jobs held by survey respondents (Figure IV.7). In Quarter 1, 71 percent of jobs in the survey met our preferred definition of a reportable job. By calendar Quarter 9, this had risen to 83 percent. Using different definitions of a reportable job and different time periods suggests that the NDNH probably did not capture anywhere from one-tenth to one-third of jobs reported on the survey.

Figure IV.7. Number of NDNH-reportable jobs



Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Note: Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey. For estimates by study group, see Appendix Table A.10.

b. Survey accuracy

Survey respondents reported holding a larger proportion of jobs reportable to the NDNH in Quarter 9 than in Quarter 1 (Figure IV.7). The average customer listed 0.35 NDNH-reportable jobs on the survey in Quarter 1. This corresponds to only 55 percent of the jobs actually captured in the NDNH for the quarter. By calendar Quarter 9, in contrast, the average number of NDNH-reportable jobs listed in the survey was 0.74, or 83 percent of jobs in the NDNH. This indicates that survey respondents were less likely to provide information on a job from Quarter 1 than they were for a job from Quarter 9. This pattern held for all definitions of reportable jobs and suggests that recall error is a potentially serious concern when using survey data alone to assess employment outcomes, especially when respondents are asked to recall jobs held more than a year in the past.

This page has been left blank for double-sided copying.

V. CONCLUSIONS

In this particular study, the main substantive conclusion was the same across data sources used: WIA-funded intensive services were effective at improving the earnings of customers. However, the magnitude of the estimated impact differed depending on the data source used to estimate it, from a high of $2,710 in Year 1 using survey data to a low of $1,271 in Year 2 using NDNH data. Either way, the services were effective.

Findings across data sources were consistent with previous studies of the differences in impacts between survey and administrative data, which have found that administrative data impacts tend to be smaller (for example, see Barnow and Greenberg 2015). This pattern held for the administrative NDNH data. Interestingly, although the overall pattern also held for administrative tax data, the differences in point estimates between administrative tax data and survey data were not nearly as large as those between NDNH and survey data. This is likely because the administrative tax data provide greater coverage. However, without linking the administrative tax data and survey data, we cannot draw more definitive conclusions. Additionally, unless we assert that the survey data are "more correct" than the NDNH data, we cannot conclude that these patterns indicate the administrative tax data are superior to the NDNH in answering research questions for evaluations of workforce development programs.

From an in-depth exploration of the differences between survey and NDNH data, we conclude that the two sources likely produce different estimates of earnings for three key reasons, all of which stem from factors that affect commonly used survey designs and administrative data sources:

1. **Survey-reported earnings in a given job exceeded those in the NDNH for the same job early in the follow-up period.** These differences disappeared by the end of the follow-up period. The method of calculating earnings based on hours worked and the wage rate was the same during the follow-up period covered, and other sources of bias such as systematic underreporting to the NDNH would not be expected to vary during the follow-up period. Thus, we conclude that overreporting of earnings in a given job on the survey seems to be the result of recall error (forgetting or misremembering) on the part of survey respondents early in the follow-up period. In addition, we do not think the method of calculating earnings is responsible for the overreporting, because people in this sample do not hold jobs for very long.[23]
2. **Survey respondents typically underreported the number of jobs they held early in the follow-up period.** Like the analysis for earnings in a given job, we found differences between the number of jobs respondents reported having on the survey and those they were shown to have using NDNH early in the reporting period. This was no longer a factor later in the follow-up period. This could again be explained by recall error when survey respondents are asked to recall jobs they held early in the follow-up period.

[23] Among those who worked in the first quarter after random assignment, only about 57 percent reported holding the same job in the fifth quarter after random assignment, and only about 35 percent reported holding the same job at the time of the 30-month follow-up survey. This suggests limited scope for this type of wage increase.

3. **Many survey respondents worked in jobs that the NDNH administrative data does not capture, such as informal work, day labor, self-employment, independent contracting, and some other alternative work arrangements.** Our estimates suggest that the NDNH was unlikely to capture one-tenth to one-third of jobs survey respondents held because it omits many informal jobs reported in the survey. Previous literature estimated rates of 9 to 14 percent of informal jobs were omitted from administrative data sources, so our results might indicate the growing economic importance of jobs not reportable to the NDNH, namely jobs in the so-called gig economy, independent contract work, self-employment, and some other alternative arrangements (Abraham et al. 2017; Katz and Krueger 2016).

To summarize, several factors seem to work in different directions at various times during the follow-up period. Survey data about jobs held early in the follow-up period are subject to recall error, and respondents are more likely to forget (or otherwise fail to report) having held some jobs in that period. Yet, for the jobs they do report having early in the follow-up period, they tend to report having earned more than the NDNH would suggest. Self-employment and contract work also tend to drive up survey-reported earnings relative to those implied by the NDNH. Because we observed that survey and NDNH earnings were roughly in line in the first four quarters after random assignment, we suggest that these forces roughly cancel each other out. In the later follow-up periods, when recall error was not as bad but people were still employed in jobs that the NDNH does not cover, we predictably found greater differences between the data sources, with survey-reported earnings quite a bit larger than NDNH. These conclusions have implications for the impact analyses examined in this report: the overall reductions in earnings levels can proportionally reduce impact estimates, even if WIA-funded services did not affect the types of jobs individuals held.

Two differences in biases across groups in this study also magnify the differences across data sources. Our results suggest that members of the core group (but not of the other study groups) had higher earnings per job in the NDNH than in the survey data late in the study period. This will attenuate the positive impact of WIA-funded intensive services (with or without training) during the later quarters examined. Additionally, our analysis suggests that survey respondents in the core-and-intensive group earned more later in the follow-up period than non-respondents. We did not find similar patterns for the core or full-WIA groups, which will further lead to estimating smaller impacts of access to intensive services, and larger impacts of access to training, when using data from the NDNH. However, these two differences in bias might be specific to the WIA Gold Standard Evaluation, so we should be cautious in concluding that similar issues would arise in other evaluations.

The evidence suggests that both surveys and administrative data have strengths and drawbacks when used to assess employment and earnings outcomes for research purposes. Because measures created from both sources differ from ideal research measures, caution should be taken when limiting a study to one source of data. Administrative data sources do not include some types of employment that are becoming important with the growth of the gig economy. Therefore, if a study population includes people likely to be highly engaged in informal labor, a survey might be more useful. Yet survey data, particularly for early in a 30-month follow-up period, suffer from recall errors that introduce unquantifiable biases—respondents are more likely to both underreport the jobs they held early in the follow-up period and overreport their

earnings in the jobs they do report having held. Therefore, for a study with a longer follow-up period, administrative data sources that contain the outcomes of interest might be the best option.

Regardless of which type of data set a study uses, it is important for researchers and consumers of research to be cognizant of data limitations, particularly when interpreting impacts derived solely from one data source.

This page has been left blank for double-sided copying.

REFERENCES

Abraham, Katharine G., John C. Haltiwanger, Kristin Sandusky, and James R. Spletzer. "Exploring Differences in Employment between Household and Establishment Data." *Journal of Labor Economics*, vol. 31, no. S1, 2013, pp. 129–172.

Abraham, Katharine G., John C. Haltiwanger, Kristin Sandusky, and James R. Spletzer. "Measuring the Gig Economy: Current Knowledge and Open Issues." *IZA Labor Statistics Workshop*: *Changing Structure of Work*. Bonn, Germany: IZA – Institute of Labor Economics, 2017.

Abraham, Katharine G., John C. Haltiwanger, Kristin Sandusky, and James R. Spletzer. "Measuring the Gig Economy: Current Knowledge and Open Issues." NBER Working Paper No. 24950. Cambridge, MA: National Bureau of Economic Research, 2018.

Andersson, Fredrik, Harry Holzer, Julia Lane, David Rosenblum, and Jeffrey Smith. "Does Federally-Funded Job Training Work? Nonexperimental Estimates of WIA Training Impacts Using Longitudinal Data on Workers and Firms." NBER Working Paper No. 19446. Cambridge, MA: National Bureau of Economic Research, 2013.

Barnow, Burt S., and David Greenberg. "Do Estimated Impacts on Earnings Depend on the Source of the Data Used to Measure Them? Evidence from Previous Social Experiments." *Evaluation Review*, vol. 39, no. 2, 2015, pp. 179–228.

Blakemore, Arthur E., Paul L. Burgess, Stuart A. Low, and Robert D. St. Louis. "Employer Tax Evasion in the Unemployment Insurance Program." *Journal of Labor Economics*, vol. 14, no. 2, 1996, pp. 210–230.

Bollinger, Christopher, and Barry Hirsch. "Is Earnings Nonresponse Ignorable?" *Review of Economics and Statistics*, vol. 95, no. 2, 2013, pp. 407–416.

Bound, John, Charles Brown, and Nancy Mathiowetz. "Measurement Error in Survey Data." Ann Arbor, MI: Population Studies Center, University of Michigan, 2000. Available at https://www.psc.isr.umich.edu/pubs/pdf/rr00-450.pdf. Accessed October 24, 2017.

Bureau of Labor Statistics. "Average Weekly Hours and Overtime of Production and Nonsupervisory Employees on Private Nonfarm Payrolls by Industry Sector, Seasonally Adjusted." Washington, DC: Bureau of Labor Statistics, 2017. Available at https://www.bls.gov/news.release/empsit.t23.htm. Accessed December 19, 2017.

Bureau of Labor Statistics. "Contingent and Alternative Employment Arrangements – May 2017." Washington, DC: Bureau of Labor Statistics, 2018. Available at https://www.bls.gov/news.release/pdf/conemp.pdf. Accessed September 13, 2018.

D'Amico, Ronald, Kate Dunham, Verenice Chavoya-Perez, Deborah Kogan, Melissa Mack, Marian Negoita, Anne Paprocki, Sheena McConnell, and Linda Rosenberg. "Providing Public Workforce Services to Job Seekers: Implementation Findings on the WIA Adult and Dislocated Worker Programs." Washington, DC: Mathematica Policy Research, 2015.

Fortson, Kenneth, Dana Rotz, Paul Burkander, Annalisa Mastri, Peter Schochet, Linda Rosenberg, Sheena McConnell, and Ronald D'Amico. "Providing Public Workforce Services to Job Seekers: 30-Month Impact Findings on the WIA Adult and Dislocated Worker Programs." Washington, DC: Mathematica Policy Research, 2017.

Groves, Robert M., and Emilia Peytcheva. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-analysis." *Public Opinion Quarterly*, vol. 72, no. 2, 2008, pp. 167–189.

Heinrich, Carolyn J., Peter Mueser, and Kenneth R. Troske. "Workforce Investment Act Non-Experimental Net Impact Evaluation: Final Report." Washington, DC: IMPAQ International, 2008.

Heinrich, Carolyn J., Peter Mueser, Kenneth R. Troske, Kyung-Seong Jeon, and Daver C. Kahvecioglu. "Do Public Employment and Training Programs Work?" *IZA Journal of Labor Economics*, vol. 2, no. 1, 2013, pp. 1–23.

Hotz, V. Joseph, and John Karl Scholz. "Measuring Employment and Income Outcomes for Low Income Populations with Administrative and Survey Data." In *Studies of Welfare Populations: Data Collection and Research Issues*, edited by Michele Ver Ploeg, Robert A. Moffitt, and Constance F. Citro. Washington, DC: National Academy Press, 2002.

Katz, Lawrence F., and Alan B. Krueger. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015." NBER Working Paper No. w22667. Cambridge, MA: National Bureau of Economic Research, 2016.

Madow, William G., Harold Nisselson, and Ingram Olkin (eds.). *Incomplete Data in Sample Surveys, Volume 1: Report and Case Studies*. New York: Academic Press, 1983a.

Madow, William G., Ingram Olkin, and Donald B. Rubin (eds.). *Incomplete Data in Sample Surveys, Volume 2: Theory and Bibliographies*. New York: Academic Press, 1983b.

Manoli, Day and Ankur, Patel. The Effects of Employment and Earnings on Tax Filing and Tax Liability: Evidence on Short-term and Long-term Effects Using Administrative Tax Data. Office of Tax Analysis, presentation slides, 2018.

Manoli, Day, Marios Michaelides, and Ankur Patel. "Long-Term Effects of Job-Search Assistance: Experimental Evidence Using Administrative Tax Data." Austin, TX: University of Texas at Austin, March 2018. Available at http://www.daymanoli.com/portfolio/nvrea/. Accessed on September 26, 2018.

Mastri, Annalisa, Sheena M. McConnell, Linda Rosenberg, Peter Z. Schochet, Dana Rotz, Andrew Clarkwest, Ken Fortson, AnnaMaria McCutcheon, Katie Bodenlos, Jessica Ziegler, and Paul Burkander. "Evaluating National Ongoing Programs: Implementing the WIA Adult and Dislocated Worker Programs Gold Standard Evaluation." Washington, DC: Mathematica Policy Research, 2015.

McConnell, Sheena M., Kenneth Fortson, Dana Rotz, Peter Z. Schochet, Paul Burkander, Linda Rosenberg, Annalisa Mastri, and Ronald D'Amico. "Do Job Seekers Benefit from Services Provided by the WIA Adult and Dislocated Worker Programs? Impacts at 15 Months." Washington, DC: Mathematica Policy Research, May 2016.

Rotz, Dana, Paul Burkander, Mary Grider, Kenneth Fortson, Linda Molinari, Elias Sanchez-Eppler, and Lindsay Cattell. "Providing Public Workforce Services to Job Seekers: 30-Month Impact Findings on the WIA Adult and Dislocated Worker Programs. Technical Supplement." Washington, DC: Mathematica Policy Research, 2017.

Schochet, Peter Z., Sheena M. McConnell, and John A. Burghardt. "National Job Corps Study: Findings Using Administrative Earnings Records Data." Princeton, NJ: Mathematica Policy Research, 2003.

Smith, Jeff. "Measuring Earnings Levels Among the Poor: Evidence from Two Samples of JTPA Eligibles." London, Ontario, Canada: University of Western Ontario, June 1997. Available at http://www-personal.umich.edu/~econjeff/Papers/Smith (1997) Measuring Earnings Levels Among the Pooer.pdf. Accessed October 19, 2017.

U.S. Department of Health and Human Services. "A Guide to the National Directory of New Hires." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families Office of Child Support Enforcement, 2017. Available at https://www.acf.hhs.gov/sites/default/files/programs/css/a_guide_to_the_national_directory_of_new_hires.pdf. Accessed October 18, 2017.

Wallace, Geoffrey L., and Robert Haveman. "The Implications of Differences Between Employer and Worker Employment/Earnings Reports for Policy Evaluation." *Journal of Policy Analysis and Management*, vol. 20, no. 4, 2007, pp. 737–754.

What Works Clearinghouse. "Assessing Attrition Bias." Washington, DC: U.S. Department of Education, Institute of Education Sciences, 2013. Available at https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_attrition_v2.1.pdf. Accessed October 18, 2017.

What Works Clearinghouse. "Assessing Attrition Bias - Addendum." Washington, DC: U.S. Department of Education, Institute of Education Sciences, 2014. Available at https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_attrition_v3.0.pdf. Accessed October 18, 2017.

This page has been left blank for double-sided copying.

APPENDIX A

ADDITIONAL TABLES

This page has been left blank for double-sided copying.

Table A.1. Survey-reported earnings, by calendar quarter versus quarter after random assignment

	Means		
	Calendar quarter (1)	Quarter after random assignment (2)	Difference (1 – 2)
Full-WIA group			
Total earnings: ($)			
Quarter 1	1,845	1,036	809*
			(0.000)
Quarter 2	2,813	1,898	914*
			(0.000)
Quarter 3	3,434	2,893	541*
			(0.000)
Quarter 4	3,632	3,516	115
			(0.250)
Quarter 5	4,229	3,390	839*
			(0.000)
Quarter 6	4,867	4,409	459*
			(0.000)
Quarter 7	4,999	4,931	68
			(0.070)
Quarter 8	5,192	5,030	163*
			(0.007)
Quarter 9	5,197	5,261	-64
			(0.192)
Quarters 1–9	36,209	32,365	3,844*
			(0.000)
Sample size	**1,588**	**1,588**	
Core-and-intensive group			
Total earnings: ($)			
Quarter 1	2,729	1,377	1,352*
			(0.000)
Quarter 2	3,463	3,047	415*
			(0.000)
Quarter 3	3,874	3,728	146*
			(0.047)
Quarter 4	4,106	3,849	257*
			(0.000)
Quarter 5	4,580	4,072	508*
			(0.000)
Quarter 6	4,899	4,760	140*
			(0.006)
Quarter 7	5,029	4,775	254*
			(0.000)
Quarter 8	5,362	5,146	216*
			(0.000)
Quarter 9	5,494	5,481	13
			(0.781)
Quarters 1–9	39,535	36,235	3,301*
			(0.000)
Sample size	**1,558**	**1,558**	

	Means		
	Calendar quarter (1)	Quarter after random assignment (2)	Difference (1 – 2)
Core group			
Total earnings: ($)			
Quarter 1	2,241	1,121	1,119* (0.000)
Quarter 2	3,020	2,451	569* (0.000)
Quarter 3	3,058	3,022	36 (0.575)
Quarter 4	3,325	3,006	319* (0.006)
Quarter 5	3,652	3,231	421* (0.000)
Quarter 6	4,277	3,790	487* (0.000)
Quarter 7	4,410	4,470	-61 (0.506)
Quarter 8	4,527	4,445	82 (0.055)
Quarter 9	4,527	4,656	-129* (0.014)
Quarters 1–9	33,037	30,193	2,844 (0.000)
Sample size	**1,553**	**1,553**	

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys.

Notes: Dollars are 2012 dollars. Estimates include zeros for those who were not employed in the corresponding time period. Estimated means and differences are regression-adjusted to account for the study design. Data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

*Significantly different from zero at the .05 level.

Table A.2. Employment by year since random assignment

	Means			
	Survey respondents in survey	All customers in NDNH	All customers in administrative tax data, W-2 data only	All customers in administrative tax data, W-2 and 1099-MISC
All study groups, weighted to be nationally representative[a]				
Employment: (%)				
Year 1	81.6	79.8	81.1	83.6
Year 2	n.a.	81.5	82.6	85.0
Sample size	**4,699**	**33,773**	**34,250**	**34,250**
Full-WIA group, weighted to be nationally representative[a]				
Employment: (%)				
Year 1	81.8	79.8	81.0	83.6
Year 2	n.a.	81.5	82.5	84.9
Sample size	**1,588**	**29,710**	**30,137**	**30,137**
Core-and-intensive group, weighted to be nationally representative[a]				
Employment: (%)				
Year 1	83.7	81.0	84.5	86.2
Year 2	n.a.	83.0	86.3	88.7
Sample size	**1,558**	**2,034**	**2,057**	**2,057**
Core group, weighted to be nationally representative[a]				
Employment: (%)				
Year 1	81.1	82.7	83.7	85.8
Year 2	n.a.	81.6	82.1	85.4
Sample size	**1,553**	**2,029**	**2,056**	**2,056**
All study groups, not weighted				
Employment: (%)				
Year 1	82.4	80.0	80.9	83.8
Year 2	86.6	81.1	82.0	84.7
Year 3	n.a.	80.8	81.6	81.7
Sample size	**4,699**	**33,773**	**34,250**	**34,250**
Full-WIA group, not weighted				
Employment: (%)				
Year 1	81.7	79.8	80.8	83.6
Year 2	87.9	80.9	81.9	84.5
Year 3	n.a.	80.6	81.5	84.3
Sample size	**1,588**	**29,710**	**30,137**	**30,137**
Core-and-intensive group, not weighted				
Employment: (%)				
Year 1	82.4	80.5	82.0	84.2
Year 2	86.3	82.0	83.8	86.2
Year 3	n.a.	81.8	83.3	85.3
Sample size	**1,558**	**2,034**	**2,057**	**2,057**

Table A.2 (*continued*)

	Means			
	Survey respondents in survey	All customers in NDNH	All customers in administrative tax data, W-2 data only	All customers in administrative tax data, W-2 and 1099-MISC
Core group, not weighted				
Employment: (%)				
Year 1	83.0	80.7	81.8	84.7
Year 2	85.7	81.5	82.9	85.8
Year 3	n.a.	81.0	81.3	84.2
Sample size	**1,553**	**2,029**	**2,056**	**2,056**

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys, Manoli and Patel (2018), and National Directory of New Hires.

Notes: Estimated means are regression-adjusted to account for the study design.

[a]Data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

n.a. = Not available.

Table A.3. Earnings by year since random assignment

	Means			
	Survey respondents in survey	All customers in NDNH	All customers in administrative tax data, W-2 data only	All customers in administrative tax data, W-2 and 1099-MISC
All study groups, weighted to be nationally representative[a]				
Total earnings: ($)				
Year 1	15,511	13,603	14,419	15,158
Year 2	n.a.	17,480	18,352	19,198
Sample size	**4,699**	**33,773**	**34,250**	**34,250**
Full-WIA group, weighted to be nationally representative[a]				
Total earnings: ($)				
Year 1	15,708	13,693	14,508	15,245
Year 2	n.a.	17,584	18,451	19,295
Sample size	**1,588**	**29,710**	**30,137**	**30,137**
Core-and-intensive group, weighted to be nationally representative[a]				
Total earnings: ($)				
Year 1	16,774	14,611	15,955	16,719
Year 2	n.a.	17,588	18,956	19,867
Sample size	**1,558**	**2,034**	**2,057**	**2,057**
Core group, weighted to be nationally representative[a]				
Total earnings: ($)				
Year 1	14,064	13,053	13,537	14,175
Year 2	n.a.	16,317	17,170	18,148
Sample size	**1,553**	**2,029**	**2,056**	**2,056**
All study groups, not weighted				
Total earnings: ($)				
Year 1	16,596	14,093	14,891	15,783
Year 2	20,409	17,938	18,910	19,902
Year 3	n.a.	20,207	21,119	22,143
Sample size	**4,699**	**33,773**	**34,250**	**34,250**
Full-WIA group, not weighted				
Total earnings: ($)				
Year 1	16,961	14,028	14,825	15,704
Year 2	20,413	17,885	18,838	19,825
Year 3	n.a.	20,222	21,055	22,094
Sample size	**1,588**	**29,710**	**30,137**	**30,137**
Core-and-intensive group, not weighted				
Total earnings: ($)				
Year 1	16,668	14,476	15,517	16,386
Year 2	21,148	18,474	19,664	20,724
Year 3	n.a.	20,498	21,988	23,003
Sample size	**1,558**	**2,034**	**2,057**	**2,057**

Table A.3 (*continued*)

	Means			
	Survey respondents in survey	All customers in NDNH	All customers in administrative tax data, W-2 data only	All customers in administrative tax data, W-2 and 1099-MISC
Core group, not weighted				
Total earnings: ($)				
Year 1	16,207	14,674	15,241	16,029
Year 2	19,838	18,187	19,223	20,048
Year 3	n.a.	20,037	21,195	22,131
Sample size	**1,553**	**2,029**	**2,056**	**2,056**

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys, Manoli and Patel (2018), and National Directory of New Hires.

Notes: Dollars are 2012 dollars. Estimates include zeros for those who were not employed in the corresponding time period. Estimated means are regression-adjusted to account for the study design.

[a]Data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

n.a. = Not available.

Table A.4. Annual impacts on employment: Unweighted data

			Administrative Tax Data	
	Survey	NDNH	W-2	W-2 and 1099-MISC
Impacts of WIA-funded training				
Employment: (percentage point impact)				
Year 1	-0.7	-0.8	-1.2	-0.6
	(0.677)	(0.436)	(0.216)	(0.456)
Year 2	1.6	-1.1	-2.0*	-1.7*
	(0.435)	(0.283)	(0.026)	(0.027)
Year 3	n.a.	-1.2	-1.7*	-0.9
		(0.222)	(0.031)	(0.183)
Impacts of WIA-funded intensive services				
Employment: (percentage point impact)				
Year 1	-0.6	-0.2	0.2	-0.6
	(0.712)	(0.908)	(0.875)	(0.611)
Year 2	0.6	0.5	1.0	0.4
	(0.727)	(0.686)	(0.384)	(0.716)
Year 3	n.a.	0.8	2.0	1.0
		(0.521)	(0.080)	(0.305)
Impacts of WIA-funded training and intensive services				
Employment: (percentage point impact)				
Year 1	-1.2	-0.9	-1.0	-1.2
	(0.388)	(0.214)	(0.189)	(0.113)
Year 2	2.2	-0.6	-1.0	-1.4
	(0.245)	(0.401)	(0.161)	(0.081)
Year 3	n.a.	-0.4	0.3	0.1
		(0.616)	(0.735)	(0.913)
Sample size	**4,699**	**33,773**	**34,250**	**34,250**

Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys, Manoli and Patel (2018), and National Directory of New Hires.

Note: Estimated impacts are regression-adjusted to account for the study design. Reported *p*-values for impacts are in parentheses and based on two-tailed t-tests.

n.a. = Not available.

*Significantly different from zero at the .05 level.

Table A.5. Annual impacts on earnings: Unweighted data

			Administrative Tax Data	
	Survey	NDNH	W-2	W-2 and 1099-MISC
Impacts of WIA-funded training				
Earnings: ($)				
Year 1	293	-447	-692	-682
	(0.619)	(0.248)	(0.059)	(0.085)
Year 2	-734	-589	-826	-898*
	(0.451)	(0.162)	(0.062)	(0.048)
Year 3	n.a.	-275	-934*	-909*
		(0.568)	(0.039)	(0.049)
Impacts of WIA-funded intensive services				
Earnings: ($)				
Year 1	462	-198	275	357
	(0.45)	(0.667)	(0.541)	(0.436)
Year 2	1,310	287	441	676
	(0.246)	(0.53)	(0.393)	(0.216)
Year 3	n.a.	461	793	871
		(0.421)	(0.241)	(0.218)
Impacts of WIA-funded training and intensive services				
Earnings: ($)				
Year 1	755	-646	-417	-325
	(0.323)	(0.215)	(0.383)	(0.531)
Year 2	575	-302	-385	-223
	(0.615)	(0.506)	(0.455)	(0.701)
Year 3	n.a.	186	-141	-38
		(0.708)	(0.814)	(0.954)
Sample size	**4,699**	**33,773**	**34,250**	**34,250**

Sources: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys, Manoli and Patel (2018), and National Directory of New Hires.

Note: Dollars are 2012 dollars. Estimates include zeroes for those who were not employed in the corresponding time period. Estimated impacts are regression-adjusted to account for the study design. Reported *p*-values for impacts are in parentheses and based on two-tailed t-tests.

n.a. = Not available.

*Significantly different from zero at the .05 level.

Table A.6. Earnings by quarter since random assignment

	Means			Conditional differences		
	All customers in NDNH (1)	Survey respondents in NDNH (2)	Survey respondents in survey (3)	1 – 2	1 – 3	2 – 3
All study groups						
Total earnings: ($)						
Quarter 1	1,906	1,681	1,880	224	26	-198*
				(0.104)	(0.649)	(0.002)
Quarter 2	2,675	2,500	2,829	175	-153	-329*
				(0.172)	(0.080)	(0.001)
Quarter 3	3,087	2,993	3,411	93	-325*	-418*
				(0.761)	(0.001)	(0.000)
Quarter 4	3,418	3,512	3,601	-94	-183	-88
				(0.727)	(0.139)	(0.713)
Quarter 5	3,702	3,464	4,161	238	-458*	-696*
				(0.511)	(0.000)	(0.000)
Quarter 6	3,962	4,068	4,786	-107	-824*	-718*
				(0.266)	(0.000)	(0.000)
Quarter 7	4,217	4,224	4,930	-7	-713*	-706*
				(0.771)	(0.000)	(0.001)
Quarter 8	4,314	4,216	5,142	97	-828*	-926*
				(0.877)	(0.000)	(0.000)
Quarter 9	4,527	4,531	5,166	-4	-639*	-635*
				(0.802)	(0.000)	(0.001)
Quarters 1–9	31,807	31,191	35,905	616	-4,098*	-4,714*
				(0.918)	(0.000)	(0.000)
Sample size	**33,773**	**4,636**	**4,699**			
Full-WIA group						
Total earnings: ($)						
Quarter 1	1,880	1,635	1,845	244	35	-210*
				(0.082)	(0.646)	(0.003)
Quarter 2	2,669	2,478	2,813	191	-144	-335*
				(0.210)	(0.172)	(0.002)
Quarter 3	3,104	3,019	3,434	86	-330*	-416*
				(0.841)	(0.003)	(0.000)
Quarter 4	3,432	3,595	3,632	-163	-200	-37
				(0.644)	(0.143)	(0.907)
Quarter 5	3,746	3,508	4,229	238	-483*	-721*
				(0.606)	(0.000)	(0.000)
Quarter 6	3,983	4,114	4,867	-131	-885*	-753*
				(0.269)	(0.000)	(0.000)
Quarter 7	4,246	4,295	4,999	-49	-754*	-705*
				(0.686)	(0.000)	(0.002)
Quarter 8	4,335	4,248	5,192	87	-857*	-944*
				(0.943)	(0.000)	(0.000)
Quarter 9	4,556	4,558	5,197	-3	-642*	-639*
				(0.820)	(0.000)	(0.002)
Quarters 1–9	31,951	31,451	36,209	500	-4,259*	-4,759*
				(0.996)	(0.000)	(0.000)
Sample size	**29,710**	**1,563**	**1,588**			
Core-and-intensive group						
Total earnings: ($)						
Quarter 1	2,564	2,591	2,729	-27	-165	-139
				(0.867)	(0.329)	(0.495)

Table A.6 (*continued*)

	Means			Conditional differences		
	All customers in NDNH (1)	Survey respondents in NDNH (2)	Survey respondents in survey (3)	1 – 2	1 – 3	2 – 3
Quarter 2	3,272	3,234	3,463	38	-191*	-228*
				(0.585)	(0.012)	(0.009)
Quarter 3	3,508	3,552	3,874	-44	-367*	-323*
				(0.908)	(0.000)	(0.000)
Quarter 4	3,562	3,631	4,106	-68	-543*	-475*
				(0.932)	(0.000)	(0.000)
Quarter 5	3,728	3,931	4,580	-203	-851*	-649*
				(0.106)	(0.000)	(0.000)
Quarter 6	4,161	4,345	4,899	-184	-738*	-554*
				(0.206)	(0.000)	(0.000)
Quarter 7	4,255	4,550	5,029	-295*	-774*	-479*
				(0.012)	(0.000)	(0.000)
Quarter 8	4,399	4,740	5,362	-341*	-962*	-622*
				(0.014)	(0.000)	(0.000)
Quarter 9	4,355	4,606	5,494	-251*	-1,139*	-888*
				(0.048)	(0.000)	(0.000)
Quarters 1–9	33,805	35,179	39,535	-1,374	-5,731*	-4,356*
				(0.219)	(0.000)	(0.000)
Sample size	**2,034**	**1,540**	**1,558**			
Core group						
Total earnings: ($)						
Quarter 1	2,275	2,116	2,241	159	34	-125
				(0.226)	(0.785)	(0.191)
Quarter 2	2,907	2,842	3,020	65	-114	-178*
				(0.730)	(0.118)	(0.025)
Quarter 3	2,915	2,838	3,058	77	-143	-220*
				(0.216)	(0.167)	(0.012)
Quarter 4	3,264	3,326	3,325	-62	-61	1
				(0.773)	(0.783)	(0.998)
Quarter 5	3,204	3,168	3,652	35	-449*	-484*
				(0.332)	(0.000)	(0.000)
Quarter 6	3,750	3,782	4,277	-33	-528*	-495*
				(0.646)	(0.000)	(0.000)
Quarter 7	4,017	4,085	4,410	-68	-393*	-325*
				(0.427)	(0.000)	(0.004)
Quarter 8	4,091	4,214	4,527	-123	-436*	-313*
				(0.210)	(0.003)	(0.023)
Quarter 9	4,338	4,391	4,527	-54	-189	-135
				(0.340)	(0.134)	(0.324)
Quarters 1–9	30,760	30,763	33,037	-4	-2,277*	-2,274*
				(0.689)	(0.011)	(0.007)
Sample size	**2,029**	**1,533**	**1,553**			

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Notes: Dollars are 2012 dollars. Estimates include zeros for those who were not employed in the corresponding time period. Estimated means and conditional differences are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

*Significantly different from zero at the .05 level.

Table A.7. Impacts on earnings by quarter since random assignment

	Means			Conditional differences		
	All customers in NDNH (1)	Survey respondents in NDNH (2)	Survey respondents in survey (3)	1 – 2	1 – 3	2 – 3
Full-WIA group vs. Core-and-intensive group contrast						
Total earnings: ($)						
Quarter 1	-683*	-933*	-865	250	182	-69
	(0.011)	(0.005)	(0.112)	(0.767)	(0.545)	(0.811)
Quarter 2	-600	-735	-647	135	46	-89
	(0.094)	(0.068)	(0.106)	(0.933)	(0.800)	(0.631)
Quarter 3	-402	-523	-461	121	59	-62
	(0.230)	(0.120)	(0.164)	(0.986)	(0.654)	(0.592)
Quarter 4	-128	-81	-500	-47	372	419
	(0.386)	(0.870)	(0.165)	(0.702)	(0.123)	(0.248)
Quarter 5	17	-439	-374	456	392	-64
	(0.916)	(0.240)	(0.286)	(0.113)	(0.121)	(0.708)
Quarter 6	-177	-221	-22	45	-155	-200
	(0.343)	(0.507)	(0.925)	(0.394)	(0.247)	(0.213)
Quarter 7	-8	-249	-25	242	17	-224
	(0.962)	(0.507)	(0.909)	(0.112)	(0.860)	(0.365)
Quarter 8	-65	-489	-157	424*	93	-332
	(0.692)	(0.132)	(0.553)	(0.025)	(0.591)	(0.100)
Quarter 9	199	-29	-274	229	473*	245
	(0.348)	(0.915)	(0.312)	(0.094)	(0.004)	(0.056)
Quarters 1–9	-1,847	-3,700	-3,325	1,854	1,479	-375
	(0.132)	(0.194)	(0.181)	(0.248)	(0.352)	(0.681)
Sample size	**33,773**	**4,636**	**4,699**			
Core-and-intensive group vs. Core group contrast						
Total earnings: ($)						
Quarter 1	287	474*	490	-186	-203	-16
	(0.123)	(0.014)	(0.148)	(0.344)	(0.445)	(0.960)
Quarter 2	363	392	443	-30	-80	-51
	(0.153)	(0.122)	(0.115)	(0.787)	(0.469)	(0.683)
Quarter 3	592	713	815	-122	-223*	-101
	(0.173)	(0.070)	(0.055)	(0.608)	(0.008)	(0.228)
Quarter 4	297	310	778*	-13	-481	-468
	(0.342)	(0.345)	(0.020)	(0.833)	(0.056)	(0.164)
Quarter 5	526*	761*	925*	-235	-398*	-163
	(0.018)	(0.016)	(0.013)	(0.087)	(0.035)	(0.336)
Quarter 6	410	563	624*	-152	-214	-62
	(0.058)	(0.066)	(0.016)	(0.625)	(0.147)	(0.679)
Quarter 7	237	466*	623*	-229	-386*	-157
	(0.291)	(0.049)	(0.016)	(0.266)	(0.014)	(0.378)
Quarter 8	310*	525*	839*	-215	-528*	-314
	(0.046)	(0.005)	(0.007)	(0.380)	(0.016)	(0.092)
Quarter 9	20	212	970*	-192	-950*	-758*
	(0.927)	(0.399)	(0.003)	(0.730)	(0.000)	(0.001)
Quarters 1–9	3,042*	4,416*	6,506*	-1,375	-3,464*	-2,089
	(0.033)	(0.012)	(0.014)	(0.498)	(0.016)	(0.061)
Sample size	**33,773**	**4,636**	**4,699**			

Table A.7 (*continued*)

	Means			Conditional differences		
	All customers in NDNH (1)	Survey respondents in NDNH (2)	Survey respondents in survey (3)	1 – 2	1 – 3	2 – 3
Full-WIA group vs. Core group contrast						
Total earnings: ($)						
Quarter 1	-396	-459	-375	64	-21	-85
	(0.229)	(0.121)	(0.145)	(0.543)	(0.908)	(0.600)
Quarter 2	-238	-343	-204	105	-34	-139
	(0.221)	(0.137)	(0.293)	(0.778)	(0.849)	(0.373)
Quarter 3	190	190	353	-1	-164	-163
	(0.208)	(0.440)	(0.186)	(0.458)	(0.427)	(0.204)
Quarter 4	168	229	278	-61	-109	-49
	(0.621)	(0.691)	(0.061)	(0.808)	(0.753)	(0.935)
Quarter 5	543*	323	550*	220	-7	-227
	(0.011)	(0.200)	(0.010)	(0.732)	(0.965)	(0.152)
Quarter 6	233	341	603*	-108	-369*	-262
	(0.144)	(0.156)	(0.023)	(0.998)	(0.019)	(0.111)
Quarter 7	229	216	598*	13	-369*	-382
	(0.273)	(0.630)	(0.041)	(0.696)	(0.018)	(0.326)
Quarter 8	245	36	681*	210	-436*	-645*
	(0.143)	(0.909)	(0.047)	(0.276)	(0.040)	(0.047)
Quarter 9	219	183	696*	36	-477*	-513
	(0.163)	(0.599)	(0.050)	(0.452)	(0.025)	(0.080)
Quarters 1–9	1,195	716	3,181	479	-1,986	-2,465
	(0.324)	(0.722)	(0.060)	(0.791)	(0.088)	(0.156)
Sample size	**33,773**	**4,636**	**4,699**			

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Notes: Dollars are 2012 dollars. Estimates include zeros for those who were not employed in the corresponding time period. Estimated means and conditional differences are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

*Significantly different from zero at the .05 level.

Table A.8. Earnings by quarter since random assignment (survey responders with the same number of jobs in the NDNH and reportable jobs in the survey)

	Means			
	All groups	Full-WIA group	Core-and-intensive group	Core group
Exactly one reportable job in quarter 1 by definition 1[a]				
Quarter 1 earnings: ($)				
NDNH	3,930	3,914	5,553	4,127
Survey[e]	4,591	4,528	6,035	4,831
Difference	-661*	-614*	-482*	-705*
	(0.000)	(0.000)	(0.003)	(0.000)
Sample size	**1,030**	**311**	**336**	**383**
Exactly one reportable job in quarter 1 by definition 2b				
Quarter 1 earnings: ($)				
NDNH	3,983	3,974	5,522	4,132
Survey[e]	4,661	4,603	6,054	4,852
Difference	-678*	-629*	-533*	-720*
	(0.000)	(0.000)	(0.001)	(0.000)
Sample size	**1,011**	**306**	**331**	**374**
Exactly one reportable job in quarter 1 by definition 3c				
Quarter 1 earnings: ($)				
NDNH	4,024	4,029	5,386	4,128
Survey[e]	4,721	4,691	5,891	4,855
Difference	-697*	-662*	-505*	-728*
	(0.000)	(0.000)	(0.002)	(0.000)
Sample size	**1,159**	**361**	**373**	**425**
Exactly one reportable job in quarter 1 by definition 4d				
Quarter 1 earnings: ($)				
NDNH	3,893	3,879	5,333	4,087
Survey[e]	4,487	4,423	5,786	4,798
Difference	-594*	-544*	-452*	-711*
	(0.000)	(0.000)	(0.003)	(0.000)
Sample size	**1,203**	**374**	**391**	**438**
One or more reportable jobs in quarter 1 by definition 1a				
Quarter 1 earnings: ($)				
NDNH	4,143	4,153	5,537	4,176
Survey[e]	4,823	4,787	6,011	4,933
Difference	-680*	-634*	-475*	-757*
	(0.000)	(0.000)	(0.002)	(0.000)
Sample size	**1,118**	**332**	**370**	**416**
One or more reportable jobs in quarter 1 by definition 2b				
Quarter 1 earnings: ($)				
NDNH	4,198	4,216	5,501	4,182
Survey[e]	4,894	4,864	6,019	4,954
Difference	-696*	-648*	-518*	-772*
	(0.000)	(0.000)	(0.001)	(0.000)
Sample size	**1,098**	**327**	**364**	**407**

Table A.8 (*continued*)

	Means			
	All groups	Full-WIA group	Core-and-intensive group	Core group
One or more reportable jobs in quarter 1 by definition 3c				
Quarter 1 earnings: ($)				
NDNH	4,167	4,192	5,316	4,160
Survey[e]	4,937	4,933	5,825	4,941
Difference	-771*	-741*	-509*	-781*
	(0.000)	(0.000)	(0.001)	(0.000)
Sample size	**1,268**	**388**	**417**	**463**
One or more reportable jobs in quarter 1 by definition 4d				
Quarter 1 earnings: ($)				
NDNH	4,043	4,053	5,260	4,104
Survey[e]	4,708	4,672	5,723	4,856
Difference	-665*	-619*	-463*	-752*
	(0.000)	(0.000)	(0.002)	(0.000)
Sample size	**1,324**	**409**	**437**	**478**
Exactly one reportable job in quarter 9 by definition 1a				
Quarter 9 earnings: ($)				
NDNH	6,695	6,695	7,474	6,881
Survey[e]	6,916	6,945	7,487	6,330
Difference	-221	-249	-13	551*
	(0.155)	(0.151)	(0.874)	(0.005)
Sample size	**2,073**	**693**	**694**	**686**
Exactly one reportable job in quarter 9 by definition 2b				
Quarter 9 earnings: ($)				
NDNH	7,249	7,275	7,852	7,290
Survey[e]	7,325	7,374	7,782	6,599
Difference	-76	-98	70	691*
	(0.655)	(0.602)	(0.442)	(0.000)
Sample size	**1,850**	**620**	**623**	**607**
Exactly one reportable job in quarter 9 by definition 3c				
Quarter 9 earnings: ($)				
NDNH	6,652	6,667	7,362	6,781
Survey[e]	6,835	6,871	7,358	6,307
Difference	-183	-204	4	474*
	(0.203)	(0.205)	(0.949)	(0.013)
Sample size	**2,165**	**728**	**724**	**713**
Exactly one reportable job in quarter 9 by definition 4d				
Quarter 9 earnings: ($)				
NDNH	6,588	6,597	7,303	6,764
Survey[e]	6,765	6,797	7,290	6,293
Difference	-177	-200	13	471*
	(0.217)	(0.213)	(0.848)	(0.013)
Sample size	**2,200**	**737**	**739**	**724**

Table A.8 (*continued*)

	Means			
	All groups	Full-WIA group	Core-and-intensive group	Core group
One or more reportable jobs in quarter 9 by definition 1a				
Quarter 9 earnings: ($)				
NDNH	6,731	6,737	7,396	6,886
Survey[e]	6,972	6,996	7,497	6,392
Difference	-241	-259	-101	494*
	(0.112)	(0.117)	(0.246)	(0.011)
Sample size	**2,256**	**749**	**761**	**746**
One or more reportable jobs in quarter 9 by definition 2b				
Quarter 9 earnings: ($)				
NDNH	7,288	7,319	7,840	7,308
Survey[e]	7,379	7,423	7,855	6,657
Difference	-91	-104	-15	650*
	(0.584)	(0.561)	(0.881)	(0.001)
Sample size	**1,960**	**660**	**658**	**642**
One or more reportable jobs in quarter 9 by definition 3c				
Quarter 9 earnings: ($)				
NDNH	6,664	6,684	7,258	6,778
Survey[e]	6,872	6,902	7,381	6,360
Difference	-208	-218	-122	419*
	(0.136)	(0.153)	(0.112)	(0.024)
Sample size	**2,364**	**789**	**799**	**776**
One or more reportable jobs in quarter 9 by definition 4d				
Quarter 9 earnings: ($)				
NDNH	6,585	6,598	7,171	6,762
Survey[e]	6,783	6,806	7,281	6,350
Difference	-197	-208	-110	412*
	(0.159)	(0.172)	(0.145)	(0.024)
Sample size	**2,419**	**806**	**821**	**792**

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Notes: Dollars are 2012 dollars. Estimated means and differences are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

[a]Jobs respondents categorized on their survey as regular jobs, excluding those described elsewhere on the survey as self-employment.

[b]Jobs respondents categorized on their survey as regular jobs that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions. Excludes all jobs described as self-employment.

[c]Jobs respondents categorized on their survey as contractor or temp positions that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions; or as regular jobs. Excludes all jobs described as self-employment. Contractor positions, or jobs with a contracting agency which may be included in the NDNH, were distinguished on the survey from independent contracting, which is excluded.

[d]Jobs respondents categorized on their survey as contractor, temp, day labor, or on-call positions that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions; or as regular jobs. Excludes all jobs described as self-employment. Contractor positions, or jobs with a contracting agency which may be included in the NDNH, were distinguished on the survey from independent contracting, which is excluded.

[e]Estimated using the mean survey-reported earnings of customers in the same ventile of survey-reported earnings.

*Significantly different from zero at the .05 level.

Table A.9. Calendar quarter 9 job characteristics (survey respondents with exactly one job in NDNH and one reportable job in calendar quarter 9[a])

	Means			Conditional differences	
	Survey earnings exceed NDNH[b] (1)	Earnings match[c] (2)	NDNH earnings exceed survey[d] (3)	1 – 2	2 – 3
All study groups					
Average hourly wages ($)	13.52	13.86	15.11	-0.34	-1.25*
				(0.664)	(0.014)
Average hours per week	42.01	40.57	42.01	1.44	-1.44
				(0.064)	(0.086)
Most frequent industries: (%)					
Healthcare and social assistance	27.1	22.0	21.0	5.1	1.0
				(0.337)	(0.867)
Manufacturing	2.8	9.0	21.6	-6.2*	-12.6*
				(0.004)	(0.021)
Accommodation and food services	5.7	10.1	6.1	-4.4	4.0
				(0.254)	(0.256)
Retail	7.5	5.9	7.2	1.6	-1.3
				(0.474)	(0.633)
Administrative, support, waste management and remediation services	5.1	10.6	3.9	-5.5	6.7*
				(0.084)	(0.030)
Most frequent occupations: (%)					
Nursing, psychiatric, and home health aides	9.2	4.1	7.6	5.1*	-3.5
				(0.025)	(0.133)
Retail sales workers	8.9	8.2	2.7	0.7	5.5
				(0.871)	(0.132)
Information and record clerks	6.8	5.3	5.5	1.5	-0.2
				(0.369)	(0.906)
Material moving workers	6.2	2.8	5.5	3.4	-2.7
				(0.249)	(0.234)
Motor vehicle operators	5.0	5.3	3.6	-0.3	1.7
				(0.853)	(0.322)
Sample size	**707**	**688**	**678**		
Full-WIA group					
Average hourly wages ($)	13.56	13.76	15.28	-0.20	-1.52*
				(0.809)	(0.008)
Average hours per week	42.18	40.84	38.71	1.34	2.13
				(0.117)	(0.064)
Most frequent industries: (%)					
Healthcare and social assistance	28.3	21.7	20.7	6.6	1.0
				(0.272)	(0.893)
Manufacturing	2.1	8.9	23.4	-6.8*	-14.5*
				(0.004)	(0.021)
Accommodation and food services	5.5	10.6	6.1	-5.1	4.5
				(0.233)	(0.274)
Retail	7.1	5.9	6.7	1.2	-0.8
				(0.665)	(0.786)
Administrative, support, waste management and remediation services	4.5	11.0	3.0	-6.5	8.0*
				(0.059)	(0.016)
Most frequent occupations: (%)					
Nursing, psychiatric, and home health aides	10.0	4.4	8.5	5.6*	-4.1
				(0.038)	(0.152)

Table A.9 (*continued*)

	Means			Conditional differences	
	Survey earnings exceed NDNH[b] (1)	Earnings match[c] (2)	NDNH earnings exceed survey[d] (3)	1 – 2	2 – 3
Retail sales workers	8.8	8.0	1.9	0.8	6.1
				(0.862)	(0.124)
Information and record clerks	6.7	4.5	4.6	2.2	-0.1
				(0.256)	(0.960)
Material moving workers	6.5	2.6	5.7	3.9	-3.1
				(0.232)	(0.235)
Motor vehicle operators	5.1	5.7	2.9	-0.6	2.8
				(0.792)	(0.224)
Sample size	**250**	**218**	**225**		
Core-and-intensive group					
Average hourly wages ($)	13.76	17.19	14.11	-3.43*	3.08*
				(0.001)	(0.005)
Average hours per week	39.41	40.22	39.32	-0.81	0.90
				(0.544)	(0.086)
Most frequent industries: (%)					
Healthcare and social assistance	25.9	26.1	18.0	-0.2	8.1
				(0.975)	(0.090)
Manufacturing	8.2	7.7	11.2	0.5	-3.5
				(0.879)	(0.245)
Accommodation and food services	4.6	3.6	13.1	1.0	-9.5*
				(0.439)	(0.004)
Retail	11.5	3.3	7.7	8.2*	-4.4*
				(0.012)	(0.013)
Administrative, support, waste management and remediation services	5.4	3.7	9.4	1.7	-5.7*
				(0.306)	(0.017)
Most frequent occupations: (%)					
Nursing, psychiatric, and home health aides	3.1	4.0	4.1	-0.9	-0.1
				(0.666)	(0.955)
Retail sales workers	9.2	5.1	3.6	4.1	1.5
				(0.219)	(0.512)
Information and record clerks	4.9	23.5	9.1	-18.6*	14.4*
				(0.001)	(0.007)
Material moving workers	2.4	3.2	4.2	-0.8	-1.0
				(0.575)	(0.512)
Motor vehicle operators	5.4	1.3	3.8	4.1	-2.5*
				(0.053)	(0.020)
Sample size	**222**	**238**	**234**		
Core group					
Average hourly wages ($)	12.90	13.32	14.01	-0.42	-0.69
				(0.416)	(0.383)
Average hours per week	40.05	37.89	36.96	2.16*	0.93
				(0.026)	(0.631)
Most frequent industries: (%)					
Healthcare and social assistance	16.5	19.8	17.6	-3.3	2.2
				(0.408)	(0.629)
Manufacturing	7.4	7.2	6.4	0.2	0.8
				(0.939)	(0.844)
Accommodation and food services	8.5	2.8	7.9	5.7*	-5.1
				(0.002)	(0.131)

Table A.9 (*continued*)

	Means			Conditional differences	
	Survey earnings exceed NDNH[b] (1)	Earnings match[c] (2)	NDNH earnings exceed survey[d] (3)	1 – 2	2 – 3
Retail	11.3	21.8	9.7	-10.5*	12.1*
				(0.021)	(0.006)
Administrative, support, waste management and remediation services	9.7	7.0	4.4	2.7	2.6
				(0.480)	(0.489)
Most frequent occupations: (%)					
Nursing, psychiatric, and home health aides	8.1	3.3	2.7	4.8	0.6
				(0.077)	(0.607)
Retail sales workers	9.3	12.5	13.1	-3.2	-0.6
				(0.349)	(0.894)
Information and record clerks	11.7	3.5	12.3	8.2*	-8.8*
				(0.009)	(0.016)
Material moving workers	4.2	2.3	1.1	1.9	1.2
				(0.216)	(0.312)
Motor vehicle operators	2.9	5.5	7.9	-2.6	-2.4
				(0.152)	(0.453)
Sample size	**235**	**232**	**219**		

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Notes: Dollars are 2012 dollars. Estimated means and conditional differences are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

[a]Jobs respondents categorized on their survey as regular jobs, excluding those described elsewhere on the survey as self-employment.

[b]Earnings captured in NDNH are more than 10 percent below the mean survey-reported earnings of customers in the same ventile of survey-reported earnings.

[c]Earnings captured in NDNH are within 10 percent of the mean survey-reported earnings of customers in the same ventile of survey-reported earnings.

[d]Earnings captured in NDNH are more than 10 percent above the mean survey-reported earnings of customers in the same ventile of survey-reported earnings.

*Significantly different from zero at the .05 level.

Table A.10. Number of reportable jobs after random assignment in quarters 1 and 9 (survey respondents)

	Average number of jobs	Difference between survey and NDNH	Share with number of NDNH jobs greater than number of survey jobs (%)	Share with number of NDNH jobs less than number of survey jobs (%)
Total quarter 1 jobs in survey	0.496	0.142	22.5	12.2
Reportable quarter 1 jobs in survey:				
Definition 1a	0.351	0.287	28.9	6.1
Definition 2b	0.340	0.298	29.5	5.8
Definition 3c	0.412	0.226	25.5	7.8
Definition 4d	0.444	0.194	23.6	8.5
Total quarter 1 jobs in NDNH	0.638	-	-	-
Total calendar quarter 9 jobs in survey	0.922	-0.024	12.8	17.5
Reportable quarter 9 jobs in survey:				
Definition 1[a]	0.749	0.148	21.7	11.7
Definition 2[b]	0.621	0.276	28.1	8.4
Definition 3[c]	0.792	0.106	18.4	12.0
Definition 4[d]	0.814	0.083	16.9	12.4
Total quarter 9 jobs in NDNH	0.898	-	-	-
Sample size	**4,636**	-	-	-

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Notes: Estimated means are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, (4) that the job-seeker consented to the study, (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

[a]Jobs respondents categorized on their survey as regular jobs, excluding those described elsewhere on the survey as self-employment.

[b]Jobs respondents categorized on their survey as regular jobs that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions. Excludes all jobs described as self-employment.

[c]Jobs respondents categorized on their survey as contractor or temp positions that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions; or as regular jobs. Excludes all jobs described as self-employment. Contractor positions, or jobs with a contracting agency which may be included in the NDNH, were distinguished on the survey from independent contracting, which is excluded.

[d]Jobs respondents categorized on their survey as contractor, temp, day labor, or on-call positions that offered health insurance, paid vacations, paid sick days, paid holidays, or pensions; or as regular jobs. Excludes all jobs described as self-employment. Contractor positions, or jobs with a contracting agency which may be included in the NDNH, were distinguished on the survey from independent contracting, which is excluded.

Table A.11. Out-of-state employment by quarter since random assignment (all customers)

	Means				Impacts		
	All groups	Full-WIA group (F)	Core-and intensive group (C&I)	Core group (C)	F – C&I	C&I – C	F – C
Employment out of state[a]: (%)							
Quarter 1	2.9	3.0	2.9	4.7	0.0	-1.7	-1.7
					(0.997)	(0.397)	(0.277)
Quarter 2	4.4	4.4	4.0	6.2	0.5	-2.3	-1.8
					(0.625)	(0.236)	(0.169)
Quarter 3	4.9	4.9	4.3	3.5	0.7	0.7	1.4
					(0.408)	(0.233)	(0.018)
Quarter 4	5.7	5.8	4.8	4.0	1.1	0.8	1.9
					(0.231)	(0.324)	(0.013)
Quarter 5	6.0	6.2	4.3	4.4	1.8	0.0	1.8
					(0.066)	(0.980)	(0.041)
Quarter 6	6.6	6.8	4.1	7.3	2.7	-3.1	-0.5
					(0.002)	(0.091)	(0.694)
Quarter 7	6.9	7.1	5.2	6.0	1.9	-0.8	1.1
					(0.059)	(0.565)	(0.074)
Quarter 8	7.1	7.4	5.2	5.3	2.2	0.0	2.1
					(0.043)	(0.959)	(0.035)
Quarter 9	7.6	7.9	6.1	4.8	1.7	1.3	3.0
					(0.192)	(0.102)	(0.009)
Quarter 10	8.0	8.2	7.0	5.7	1.2	1.3	2.5
					(0.442)	(0.133)	(0.046)
Quarter 11	7.8	8.0	8.1	6.0	-0.2	2.2	2.0
					(0.842)	(0.076)	(0.068)
Quarter 12	8.2	8.4	7.8	7.0	0.6	0.7	1.3
					(0.480)	(0.610)	(0.321)
Employed out of state at any time in quarters 1–12	16.5	16.8	14.7	16.3	2.2	-1.6	0.5
					(0.099)	(0.494)	(0.700)
Number of quarters employed out of state	0.761	0.780	0.639	0.649	0.141	-0.009	0.132
					(0.134)	(0.909)	(0.007)
Sample size	**33,773**	**29,710**	**2,034**	**2,029**			

Source: National Directory of New Hires.

Notes: Estimated means and impacts are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study.

[a]Out-of-state employment includes individuals with any NDNH earnings reported in a state other than the one in which the LWIA where they were recruited for the WIA Gold Standard Evaluation is located.

Table A.12. Employment by quarter since random assignment

	Means			Conditional differences		
	All customers in NDNH (1)	Survey respondents in NDNH (2)	Survey respondents in survey (3)	1 – 2	1 – 3	2 – 3
All study groups						
Employment: (%)						
Quarter 1	52.6	50.3	43.6	-0.2	6.5*	6.7*
				(0.613)	(0.000)	(0.000)
Quarter 2	57.0	55.1	56.7	1.4	-0.3	-1.6
				(0.244)	(0.877)	(0.281)
Quarter 3	60.8	59.2	62.1	0.8	-2.0	-2.9*
				(0.592)	(0.182)	(0.046)
Quarter 4	62.6	62.2	67.1	0.8	-4.1*	-4.9*
				(0.561)	(0.005)	(0.014)
Quarter 5	65.8	66.2	74.1	-1.4	-9.2*	-7.9*
				(0.416)	(0.000)	(0.000)
Quarter 6	69.3	69.3	74.8	-1.9	-7.5*	-5.6*
				(0.223)	(0.000)	(0.000)
Quarter 7	68.5	68.9	76.1	-0.5	-7.8*	-7.2*
				(0.647)	(0.000)	(0.001)
Quarter 8	70.6	71.0	79.2	-2.2	-10.3*	-8.1*
				(0.071)	(0.000)	(0.000)
Quarter 9	72.4	72.3	79.5	-2.3	-9.4*	-7.1*
				(0.230)	(0.000)	(0.001)
Quarters 1–9	90.6	90.7	92.1	-1.8	-3.3*	-1.5
				(0.095)	(0.000)	(0.057)
Sample size	**33,773**	**4,636**	**4,636**			
Full-WIA group						
Employment: (%)						
Quarter 1	52.3	49.7	43.0	0.1	6.8*	6.8*
				(0.484)	(0.000)	(0.000)
Quarter 2	56.8	54.8	56.8	1.5	-0.5	-2.0
				(0.258)	(0.802)	(0.233)
Quarter 3	60.8	59.2	62.4	0.8	-2.3	-3.1*
				(0.591)	(0.151)	(0.036)
Quarter 4	63.0	62.7	67.7	0.5	-4.5*	-5.0*
				(0.758)	(0.005)	(0.021)
Quarter 5	66.2	66.6	74.6	-1.6	-9.5*	-8.0*
				(0.427)	(0.000)	(0.000)
Quarter 6	69.6	69.3	75.1	-1.8	-7.6*	-5.7*
				(0.344)	(0.000)	(0.000)
Quarter 7	68.8	69.1	76.3	-0.7	-7.8*	-7.2*
				(0.681)	(0.000)	(0.002)
Quarter 8	70.8	71.2	79.5	-2.4	-10.7*	-8.3*
				(0.103)	(0.000)	(0.001)
Quarter 9	72.7	72.3	79.7	-2.2	-9.6*	-7.4*
				(0.349)	(0.000)	(0.001)
Quarters 1–9	90.7	90.9	92.2	-2.2	-3.5*	-1.4
				(0.114)	(0.000)	(0.109)
Sample size	**29,710**	**1,563**	**1,563**			
Core-and-intensive group						
Employment: (%)						
Quarter 1	54.7	53.8	50.9	0.9	3.7	2.9
				(0.479)	(0.161)	(0.360)

Table A.12 (*continued*)

	Means			Conditional differences		
	All customers in NDNH (1)	Survey respondents in NDNH (2)	Survey respondents in survey (3)	1 – 2	1 – 3	2 – 3
Quarter 2	61.4	61.0	58.1	0.4	3.3*	2.9*
				(0.784)	(0.020)	(0.015)
Quarter 3	63.8	63.7	63.2	0.1	0.6	0.4
				(0.825)	(0.730)	(0.791)
Quarter 4	63.3	63.3	67.0	0.0	-3.7*	-3.7
				(0.809)	(0.013)	(0.058)
Quarter 5	67.3	67.3	73.1	0.0	-5.8*	-5.8*
				(0.817)	(0.010)	(0.005)
Quarter 6	69.3	71.4	75.8	-2.1	-6.5*	-4.4*
				(0.126)	(0.007)	(0.008)
Quarter 7	69.4	71.8	76.3	-2.4*	-6.9*	-4.5*
				(0.007)	(0.000)	(0.008)
Quarter 8	70.2	72.1	77.6	-2.0*	-7.4*	-5.4*
				(0.020)	(0.000)	(0.000)
Quarter 9	70.8	72.5	77.8	-1.7*	-7.1*	-5.4*
				(0.006)	(0.000)	(0.004)
Quarters 1–9	89.1	89.3	92.3	-0.2	-3.1*	-3.0*
				(0.442)	(0.000)	(0.001)
Sample size	**2,034**	**1,540**	**1,540**			
Core group						
Employment: (%)						
Quarter 1	55.3	54.8	49.9	0.5	5.4*	4.8*
				(0.452)	(0.002)	(0.000)
Quarter 2	58.9	58.1	57.5	0.8	1.3	0.6
				(0.598)	(0.349)	(0.567)
Quarter 3	62.4	62.2	60.1	0.2	2.3	2.1
				(0.976)	(0.161)	(0.120)
Quarter 4	61.7	61.4	60.4	0.3	1.3	1.0
				(0.586)	(0.453)	(0.558)
Quarter 5	62.2	62.7	67.3	-0.6	-5.1*	-4.6*
				(0.296)	(0.000)	(0.000)
Quarter 6	66.7	68.8	72.3	-2.0	-5.6*	-3.5*
				(0.069)	(0.000)	(0.001)
Quarter 7	68.5	70.2	74.2	-1.7	-5.7*	-4.0*
				(0.183)	(0.002)	(0.006)
Quarter 8	68.5	70.6	74.6	-2.1	-6.1*	-4.1
				(0.172)	(0.024)	(0.054)
Quarter 9	68.5	71.4	75.6	-2.9	-7.2*	-4.3*
				(0.051)	(0.002)	(0.027)
Quarters 1–9	91.1	89.8	89.3	1.3*	1.8	0.5
				(0.009)	(0.283)	(0.746)
Sample size	**2,029**	**1,533**	**1,533**			

Source: WIA Gold Standard Evaluation 15- and 30-month follow-up surveys and National Directory of New Hires.

Notes: Estimated means and conditional differences are regression-adjusted to account for the study design. All data are weighted to account for the probability (1) that the local area was selected to participate in the study, (2) that the local area agreed to participate in the study, (3) of assignment to each study group, and (4) that the job-seeker consented to the study. Estimates for survey respondents are additionally weighted to account for the probability (5) that the job-seeker was selected for the survey, and (6) that the job-seeker completed the survey.

*Significantly different from zero at the .05 level.

This page has been left blank for double-sided copying.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ SEATTLE, WA ■ TUCSON, AZ ■ WASHINGTON, DC ■ WOODLAWN, MD