Interactions of Public Paratransit and Vocational Rehabilitation

Working Paper 2017-12

CHRIS CLAPP, STEVEN STERN AND DAN YU

November, 2017
Interactions of Public Paratransit and Vocational Rehabilitation

Christopher M. Clapp  Steven Stern
Florida State University  Stony Brook University
Dan Yu*
University of Minnesota
November 2017

Abstract
Federal and state governments spend over $3 billion annually on public-sector Vocational Rehabilitation (VR) programs, yet almost a third of people with disabilities report having inadequate access to the transportation necessary to commute to a job, potentially negating the positive effects of these interventions. We examine this previously understudied connection by assessing the impact access to public paratransit has on measures of VR program effectiveness. To do so, we use the data and estimates from three previously estimated structural models of VR service receipt and labor market outcomes that contain limited information about mobility. We spatially link the generalized residuals from these models to different measures of the availability and efficiency of local paratransit systems to determine whether paratransit explains any of the residual variation in the short- or long-run labor market outcomes of individuals receiving VR services. Results show that access to paratransit is an important determinant of the efficacy of VR services, but that effects are heterogeneous across disability groups. We discuss the policy implications of our findings for VR programs.

1 Introduction
State vocational rehabilitation (VR) agencies provide services to people with heterogeneous disabilities. For many clients, those same conditions also create barriers to mobility. It has long-been recognized by those in the VR field that,

*We would like to thank Bob Schmidt for suggesting the approach to interpreting maintenance services, John Pepper, Anastasia Semykina, and Emek Basker for valuable comments, and numerous employees at paratransit agencies and independent living centers across the state of Virginia for their help in making sense of our transportation data. All errors are ours.
in addition to the support provided by VR agencies, quality transportation is a prerequisite for employment (e.g., West et al., 1998; Arkansas RTC, 1992; Schmidt and Smith, 2007; Magill-Evans et al., 2008; Sabella, Bezyak, and Gattis, 2016). To that end, there were 62 federal programs that spent more than $2.4 billion to fund transportation services for people with mobility issues in fiscal year 2001 (U.S. GAO, 2003). This compares to $3.2 billion spent on VR programs in the same year (U.S. Department of Education, 2001). Yet the state of the research on the connection between these two important types of public programs is incomplete.

Recent work by Dean et al. (2015, 2017a, 2017b) (DPSS) makes progress in addressing issues related to VR agency interventions. DPSS develop and implement a methodology to evaluate the impact of VR service choices on the employment and earnings outcomes of individuals with cognitive impairments, mental illness, and physical disabilities. DPSS advance the existing literature by developing a structural model that addresses the selection concerns that stem from the individualized nature of VR service receipt. While their models control for a limited set of measures of mobility, their focus is on the effect of service receipt on the labor market outcomes of VR clients. Additionally, several of their results suggest that the broad, self-reported transportation measures they include provide an incomplete accounting of the transportation options available to the individuals in their data. Particularly, DPSS include no direct measures of an important form of transportation specifically operated for individuals with disabilities: paratransit.

For disabled people who do not have personal transportation available, public paratransit is frequently their only other means of getting around, but there is ample evidence of user frustration with the quality of specialized transportation (e.g., Weller, 1994; West et al., 1998; National Organization on Disability, 2000; National Council on Disability, 2005; Jolly, Priestley, and Matthews, 2006; Schmidt and Smith, 2007; Magill-Evans et al., 2008; Denson, 2015). Additionally, the availability of these systems varies widely: the national proportion of counties providing demand-response service in 2013 was 79% (Mattson, 2015). This means that a non-trivial fraction of individuals with disabilities may receive extensive job training services from one set of federal and state agencies but have no way to commute to a job because of the lack of effective transportation services provided by another set of agencies. If so, the availability and

---

1 States and local governments spend several hundred million more dollars, but exact amounts are not known (U.S. GAO, 2003).
2 In 2000, the Federal Transit Administration reported that 73 million demand-response rides were provided across the nation (Koffman, Weiner, and Raphael, 2003).
3 Common issues affecting perceived quality of public and/or specialized transportation are accessibility, reliability, and cost (National Organization on Disability, 2000; Scheer et al., 2003; National Council on Disability, 2005).
4 Mattson (2015) reports that, in Virginia, the state this study focuses on, only 57 out of 95 counties (58%) provided those services in 2013. The discrepancy between Virginia and the national average may be overstated because the Virginia numbers exclude independent cities which are much more likely to provide services.
5 Ridership on fixed-route public transit and paratransit systems increased dramatically from 1984-1995 (Bearse et al., 2004) and then again from 1995-2005 (National Council on
quality of paratransit may play a large role in determining the labor market outcomes of VR clients.\(^6\)

However, just providing transportation may not be sufficient to get disabled people to work. Rosenbloom (2007) argues that

“perhaps the most intractable issue in current debates is the tendency of those in every other substantive field from education to employment or from recreation to health care to assume that transportation deficiencies account for all or most of the underutilization of public and private services considered essential to the well-being of those with disabilities. In fact, substantial research shows that most people with disabilities face multiple barriers to both their mobility and their ability to get an education or a job or to access a range of public and private services from grocery stores to medical facilities. The causes of and solutions to these problems are complex; policy analysts must understand and address them in sophisticated ways that extend beyond public transit networks and, indeed, beyond transportation systems alone.”

Sabella, Bezyak, and Gattis (2016) make similar statements. If they are correct, then the availability of paratransit services alone is unlikely to affect labor market outcomes in a meaningful way.

In this paper, we combine imperfect information about the availability of specialized transportation for people with disabilities with the results in DPSS to assess the impact access to and the efficacy of public paratransit has on VR program effectiveness. The overarching methodological strategy is to use the data, estimation methodology, and estimates in DPSS in concert with imperfect data on different measures of paratransit presence and usefulness that we collect. We create a novel dataset by spatially linking the generalized residuals from the previously estimated structural models of service receipt and labor market outcomes to the characteristics of paratransit systems in each individual's community. The DPSS generalized residuals (Gourieroux et al., 1987) are associated with models estimated using 10 years of merged administrative records from the Virginia Department of Aging and Rehabilitative Services (DARS) and the Virginia Employment Commission (VEC). DPSS track cohorts of individuals receiving VR services from DARS in fiscal year 2000. We augment this information with paratransit data from the U.S. Department of Transportation for those same years to create a panel dataset covering 1997-2007. We use these data to estimate correlations between measures of the extent and efficacy of local paratransit and employment generalized residuals and conditional log quarterly earnings generalized residuals. We then perform a series of pseudo-Lagrange Multiplier tests (Checkovich and Stern, 2002; Friedberg and Stern, 2005).

\(^6\) Jolly, Priestley, and Matthews (2006) find that people with disabilities are twice as likely as people without disabilities to turn down a job because of lack of transportation in Great Britain.
2014; Dean et al., 2017a) to measure the statistical significance of the paratransit variables in explaining some of the residual variation in the labor market outcomes. Given that we use data on only one cohort of individuals through time, we compare correlations for a “beneficiary group” of individuals who do not have other means of transportation and a “placebo group” of individuals who do not have a need for paratransit services. This allows us to control for macroeconomic factors affecting both groups and identify whether paratransit is important in improving labor market outcomes of VR recipients. We base these groups on two different definitions of people who are likely/unlikely to benefit from paratransit services.

We make use of a novel methodology for two reasons. First, by building from the DPSS generalized residuals, we implicitly deal with the selection concerns that their model goes to great lengths to address. By testing whether paratransit measures explain residual variation, we do so in a way that is both intuitive and computationally tractable. Intuitively, we net out the structure, control variables, and other complications from DPSS using techniques akin to those frequently used in multiple regression contexts: partialling or projecting out the effects of control variables. We then use triple difference techniques to evaluate the importance of paratransit measures in explaining the net variation relative to an unaffected baseline.

Second, because of uncertainty about the appropriate measures of available paratransit and because of the existence of noise in our measures of availability, we prefer testing over re-estimation of a DPSS-style model including some arbitrarily chosen measure of paratransit availability. One of the points of the paper is to advertise some of the benefits of pseudo-Lagrange Multiplier test statistics that work well in such a situation. Specifically, conditional on already having estimated the models in DPSS, the low computational cost of our methodology allows us to easily experiment with multiple measures of the efficacy of paratransit.\(^7\) Additionally, we are able to conduct a sensitivity analysis that indicates that our findings are somewhat robust to potential measurement error in our paratransit variables. Such an analysis would not be feasible if we were to re-estimate the complex DPSS models for each potential measure of paratransit availability and each disability group. Alternatively, we could estimate a simpler version of the DPSS models, but doing so would involve a difficult trade-off. Dean et al. (2017a) compare estimates from their structural model to those from less complex (probit and linear model) analogs. While estimates based on the two approaches are qualitatively similar, there are some key differences that lead DPSS to conclude that the modelling exercise is highly valuable.

Our results provide evidence that access to well-functioning paratransit is an important determinant of the efficacy of VR services. We separately test for the effectiveness of paratransit on employment and conditional log quarterly earnings in both the short and long run. We do so for individuals with the three different categories of disabilities analyzed in DPSS. Based on our preferred definitions of “beneficiary” and “placebo” groups, for VR clients with

\(^7\)Of course, one should be aware of possible post-hoc testing bias.
cognitive impairments or mental illness, we find evidence that higher quality
public paratransit increases both the long-run probability of being employed
and earnings conditional on being employed. For clients with mental illness,
we also find evidence of positive, significant effects in the short run. Our most
robust findings are for individuals with physical impairments. For those VR
clients, both the existence and quality of paratransit is associated with signifi-
cant increases in employment and significant decreases in conditional earnings
in both the short and long run. This suggests that paratransit may open up
employment opportunities for physically disabled individuals who would other-
wise be unable to work despite assistance from the VR agency and may cause an
individual to accept jobs with lower pay because transportation costs are lower.
To determine whether our results are economically significant, we compare their
magnitudes to those of personal the transportation effects from DPSS. In doing
so, we find paratransit effects that are of the same order of magnitude as the
impact that access to personal transportation has on the employment outcomes
of VR clients. While the patterns of employment outcome effects are similar to
those found in DPSS as a whole, these patterns vary across groups and suggest
heterogeneous mechanisms are at work. We are unable to speak to the causes
of these differences directly, but further exploration of this intriguing variation
is ripe for future research. Overall, our results suggest that a well-functioning
paratransit system can augment the work done by VR agencies and improve the
labor market outcomes of their clients.

Our work makes several important contributions. First, we address what we
feel is a significant gap in the existing literature, as we know of no research that
directly looks at the connection between vocational rehabilitation and trans-
portation programs in improving outcomes for people with disabilities. Rosen-
bloom (2007) and Sabella, Bezyak, and Gattis (2016) point out the importance
of this connection for people with disabilities. We are able to productively add
to this discussion.

Putting aside the lack of research on the effects of multiple relevant pro-
grams (e.g., VR and transportation), the analysis of the effects of paratransit
systems on labor market outcomes is thin. The majority of previous work on this
topic has focused on demand for paratransit services (Stern, 1993; Fitzgerald
et al., 2000; Bearse et al., 2004; Goodwill and Joslin, 2013; Deka and Gonza-
les, 2014) or conducted cost-benefit analysis of paratransit systems (Nguyen-
Hoang and Yeung, 2010). What little work exists on the impact paratransit has
on labor market outcomes uses elementary methods. For instance, CES, Inc.
and TranSystems (2009) (CT) evaluate the Job Access and Reverse Commute
(JARC) and New Freedom (NF) programs for the Federal Transit Administra-
ation. Rather than basing their evaluations of these programs on the impact
they had on actual outcomes relevant to the targeted population, CT provides

8The JARC and NF programs provide grants to assist communities with developing or
expanding their public transportation programs. The JARC program’s goal is to address
the transportation issues facing welfare recipients and low-income individuals, with a focus on job
access in suburban areas. The NF program specifically seeks to improve the mobility options
available to disabled people.
summary statistics of the expansion in system capacity, coverage, and usage attributable to the programs. At best, the evaluation of the JARC program provides a measure of job access, but this is calculated by determining the number of jobs per mile for a sample of routes, then multiplying this factor by the total number of new route miles. The analysis ignores obvious selection issues which make using this information to inform future grant decisions problematic. These selection problems are possibly quite severe; at the time of the analysis, less than 10% of the available funding for the NF program had been awarded (CTAA, 2016).9

Second, our work to address how two disparate public programs affect a vulnerable population has potential policy implications for both VR and paratransit decision makers. Our analysis is especially salient given the state of both programs. With respect to VR agencies, the Government Accountability Office (2005, 2012) recently emphasized the need to improve the methods used to evaluate VR service provision, and the 2014 Workforce Innovation and Opportunity Act requires agencies to report the employment and earnings outcomes of their clients as a condition of their funding (Dean et al., 2015). Information about the extent to which transportation issues limit the efficacy of VR services can help agencies improve how they allocate resources. With respect to paratransit programs, while Nguyen-Hoang and Yeung (2010) find that the aggregate benefits of paratransit exceed the costs, Rosenbloom (2007) reports that the provision of paratransit by public transportation agencies is extremely expensive, particularly in larger and lower density areas. This is due to the need for “on-demand” paratransit services that limit opportunities for economies of scale. As such, any benefits that paratransit provides to VR agencies may be useful in justifying the cost of continued operation or expansion of individual systems, especially those in currently under-served areas.

Finally, we start from the structural models of endogenous service provision and labor market outcomes in DPSS that address selection issues that were mostly previously ignored in the literature. By doing so, we are able to interpret our estimates as causal measures of the impact of paratransit on the employment and earnings of individuals who received VR services. We produce such estimates in both the short and long run for numerous measures of paratransit provision in local communities.

The remainder of the paper proceeds as follows: Sections 2 and 3 present a basic version of the models used by DPSS and their estimation strategy. More details for these steps are available in DPSS. Section 4 describes how we conduct our pseudo-Lagrange Multiplier tests. Section 5 discusses the data construction process and provides descriptive statistics about the data. Section 6 presents

9 Another potential source of data is the Bureau of Transportation Statistics Omnibus Survey (2002) with a cross-section including information on disability, available transportation, and employment. The dataset has 5019 individuals with 5% receiving special education (which is the closest proxy in the data to having a cognitive impairment), 27% having a condition that limits physical activity, and no information on mental illness prevalence. This turns out not to be large enough to say much about paratransit because less than 1% of people in the sample use paratransit as their most frequent source of transportation.
our results. The final section concludes.

2 Model

The basic model has three equations of interest: a service receipt equation, an employment equation, and a conditional earnings equation.\textsuperscript{10} Let $y_{ij}^*$ be the (latent) value to client $i$ of receiving service $j$, $j = 1, 2, \ldots, J$. There are $J = 6$ available services.\textsuperscript{11} We assume that

$$y_{ij}^* = X_i \gamma_j^y + u_{ij}^y + \varepsilon_{ij}^y$$

(1)

where $X_i$ is a vector of explanatory variables described in Section 5.1, $u_{ij}^y = (u_{i1}^y, u_{i2}^y, \ldots, u_{ij}^y)'$ is a vector of unobserved heterogeneity errors with a complex joint distribution function described in DPSS, and $\varepsilon_{ij}^y \sim iidG_y(\cdot)$ is a person/choice-specific (idiosyncratic) error. Let

$$y_{ij} = 1 (y_{ij}^* > 0)$$

(2)

be an indicator for receipt of service $j$ by $i$. Equations (1) and (2) constitute a multivariate binary discrete choice model (e.g., Greene, 2009). Each individual $i$ can choose multiple services; there are $2^6 = 64$ different service combinations available.

Next, let

$$e_{it}^* = Z_{it}\gamma_e + \sum_{j=1}^{J} y_{ij} \sum_{\tau=1}^{4} 1 (t \in T_{\tau i}) \alpha_{j\tau} + u_{it}^e + \varepsilon_{it}^e$$

(3)

be the (latent) value of working where $Z_{it}$ is a vector of explanatory variables similar to $X_i$ but possibly time-varying, $u_{it}^e$ is a vector of unobserved heterogeneity errors with a complex joint distribution function described in DPSS, and $\varepsilon_{it}^e \sim iidG_e(\cdot)$ is a person/time-specific (idiosyncratic) error. Define $t_{si}$ as the period where $i$ receives service.\textsuperscript{12} Then $(T_{1i}, T_{2i}, T_{3i}, T_{4i})$ is a partition of the periods $i$’s earnings history is observed, excluding $t_{si}$, with

- $T_{1i} = \{t : t < t_{si} - 1\}$ Pre-service;
- $T_{2i} = \{t_{si} - 1\}$ Ashenfelter dip (1978);
- $T_{3i} = \{t : t_{si} < t < t_{si} + 8\}$ Post-service short run;
- $T_{4i} = \{t : t > t_{si} + 8\}$ Post-service long run.

Allowing the effect of service receipt on $e_{it}^*$ to differ across these time segments allows us to measure the short-run effect $(\alpha_{j3} - \alpha_{j1})$ and long-run effect

\textsuperscript{10}Each of the models has some complications that are ignored here. For example, Dean et al. (2017a) includes a DI/SSI receipt equation.

\textsuperscript{11}The available services are diagnosis & evaluation, training, education, restoration, maintenance, and other. These are described in DPSS. Note that, throughout the paper, variable names like service types are put in a special font to avoid confusion.

\textsuperscript{12}We assume away multiple periods of service.
($\alpha_{j4} - \alpha_{j1})$ of each service $j$. Let
\[ e_{it} = 1 (e_{it} > 0) \]
be an indicator for $i$ working in period $t$.

Finally, let
\[ w_{it} = Z_{it} \gamma_W + \sum_{j=1}^J y_{ij} \sum_{\tau=1}^4 1 (t \in T_{\tau}) \alpha_{j\tau}^{w} + u_{it}^w + \varepsilon_{it}^w \]
be log earnings in period $t$, conditional on $e_{it} = 1$, with parallel structure to equation (3), and let $\varepsilon_{it}^w \sim iidG_w(\cdot)$ with density $g_w(\cdot)$. Equations (1) through (5), along with the error structure defined in DPSS, constitute the model.

3 Estimation

The method of estimation is maximum simulated likelihood (e.g., Börsch-Supan and Hajivassiliou, 1992; Stern, 1997). The complex structure of the errors causes the likelihood function to be a high-dimensional integral, and simulation is the obvious way to approximate the integral. Define
\[
Q^y_{ij} (u_{ij}^y, \theta) = -X_{ij} \gamma^y - u_{ij}^y,
\]
\[
Q^e_{ij} (u_{ij}^e, \theta) = -Z_{it} \gamma^e - \sum_{j=1}^J y_{ij} \sum_{\tau=1}^4 1 (t \in T_{\tau}) \alpha_{j\tau}^e - u_{it}^e,
\]
\[
Q^w_{it} (u_{it}^w, \theta) = w_{it} - Z_{it} \gamma^w - \sum_{j=1}^J y_{ij} \sum_{\tau=1}^4 1 (t \in T_{\tau}) \alpha_{j\tau}^w - u_{it}^w.
\]
The likelihood contribution for observation $i$ is
\[
L_i (\theta) = \int \left[ \prod_{j=1}^J L^y_{ij} (u_{ij}^y, \theta) \right] \left[ \prod_t L^e_{it} (u_{it}^e, u_{it}^w, \theta) \right] dF (u_i, \theta)
\]
where $\theta$ is the vector of parameters to estimate, $u_i \sim iidF(\cdot, \cdot)$ is the vector of errors for $i$,$L^y_{ij} (u_{ij}^y, \theta) = [1 - G_y (Q^y_{ij} (u_{ij}^y, \theta))]^{y_{ij}} G_y (Q^y_{ij} (u_{ij}^y, \theta))^{1-y_{ij}}$ is the likelihood component for receipt (or not) of service $j$, and
\[
L^e_{it} (u_{it}^e, u_{it}^w, \theta) = \begin{cases} G_e (Q^e_{it} (u_{it}^e, \theta)) & \text{if } e_{it} = 0 \\ [1 - G_e (Q^e_{it} (u_{it}^e, \theta))] g_w (Q^w_{it} (u_{it}^w, \theta)) & \text{if } e_{it} = 1 \end{cases}
\]
is the likelihood component for labor market outcomes, employment and quarterly earnings. The simulator of $L_i (\theta)$ is
\[
\lim_{R \to \infty} L_i (\theta) = \frac{1}{R} \sum_{r=1}^R \left[ \prod_{j=1}^J L^y_{ij} (u_{ij}^y, \theta) \right] \left[ \prod_t L^e_{it} (u_{it}^e, u_{it}^w, \theta) \right]
\]
where \( u_i = \{ u_i^j \}_{j=1}^J, \{ u_i^w \}_{j=1}^4 \) is a pseudo-random draw from \( F(\cdot, \theta) \). In practice, we also use antithetic acceleration (Geweke, 1988) to reduce simulation variance.

4 Testing

4.1 Pseudo-Lagrange Multiplier Tests

Consider a generalization of equations (3) and (5) of the form,

\[
e^{e}_{it} = Z_{it} \gamma^{e} + \sum_{j=1}^{J} y_{ij} \sum_{\tau=1}^{4} 1(t \in T_{\tau i}) (\alpha^{e}_{j\tau} + \lambda^{e}_{j\tau} b_{it}) + u^{e}_{it} + \epsilon^{e}_{it}, \quad (6)
\]

\[
w^{w}_{it} = Z_{it} \gamma^{w} + \sum_{j=1}^{J} y_{ij} \sum_{\tau=1}^{4} 1(t \in T_{\tau i}) (\alpha^{w}_{j\tau} + \lambda^{w}_{j\tau} b_{it}) + u^{w}_{it} + \epsilon^{w}_{it} \quad (7)
\]

where \( b_{it} \) is a measure of the quality of public transportation in \( i \)'s county in period \( t \) (defined in Section 5.3). Define

\[
\lambda_j = \{ \lambda^{e}_{j\tau}, \lambda^{w}_{j\tau} \}_{\tau=1}^{4},
\]

\( \lambda_{js} \) as a subset of \( \lambda_j \) of some interest, and \( \lambda_{jc} = \lambda_j \setminus \lambda_{js} \) as the complement of \( \lambda_{js} \). We want to test

\[
H_0 : \lambda_j = 0 \text{ vs } H_A : \lambda_{jc} = 0, \lambda_{js} \neq 0.
\]

For some problems, we can define \( \lambda = \{ \lambda_j \}_{j=1}^{J} \) and decompose \( \lambda \) into \( \lambda_e \) and \( \lambda_c = \lambda \setminus \lambda_e \). Note that, under \( H_0 \) and appropriate specification of \( (\lambda_{js}, \lambda_{jc}) \), equations (6) and (7) simplify to equations (3) and (5). Using a Lagrange Multiplier (LM) test is a good choice of a test statistic because it does not require re-estimating the model replacing equations (3) and (5) with equations (6) and (7). Instead, using an LM test requires computing \( \partial e^{e}_{it} / \partial \lambda_{js} \) and \( \partial w^{w}_{it} / \partial \lambda_{js} \) (which are both the appropriate vectors of the relevant paratransit data) and then using them to construct score statistics.

An even simpler approach, aggregating over services, is to simulate the generalized residuals for equations (3) and (5)\(^{13}\) and then compute the correlation of the generalized residuals and each public transportation variable of interest. If the generalized residuals for a particular labor market outcome are correlated with a specific measure of transportation quality, then the variation in transportation quality helps “explain” some of the variation in labor market

\(^{13}\)The generalized residual for equation (3) is the integral of the inverse Mills ratio over the density of the unobserved heterogeneity if we assume that \( G_e(\cdot) \) is normal. The generalized residual for equation (5) is just the OLS residual.
outcomes not “explained” by the included covariates.\footnote{In this case, one might want to focus on just positive derivatives.} Let $\hat{\rho}_{kn}^k$ be the sample correlation for $k = e, w$, constructed with $n$ observations.\footnote{We also could do this separately for each service as above. However, in Section 6, we focus on aggregated results because a) there otherwise would be too many results to present and b) there is no obvious reason for public transportation effects to vary across services other than maintenance services.} Then, under $H_0$, 

$$
\hat{\rho}_{kn}^k \left\{ \frac{n - 2}{1 - (\hat{\rho}_{kn}^k)^2} \right\} \sim N(0, 1)
$$

for large $n$. Figure 1 shows how the test statistic varies for three relatively small values of $\hat{\rho}_{kn}^k$ as $n$ gets large. Note that, for employment, $n \approx 32000$, and, for log earnings, $n \approx 10000$. Checkovitch and Stern (2002) call this test statistic a pseudo-LM test statistic. It is not an LM test statistic because it is not using the score statistics. But it is like an LM test statistic because a) there is no need to re-optimize the likelihood function and b) one is measuring how far from zero an object of interest is at a place of interest under $H_0$.

The parameter functions of most interest are $\{\lambda_{j3} - \lambda_{j1}\}_{j=1}^J$ (short-run effect) and $\{\lambda_{j4} - \lambda_{j1}\}_{j=1}^J$ (long-run effect). Define $\hat{\eta}_{it}^k$ as the generalized residual of type $k = e, w$ for $i$ in period $t$. Then the test statistics of interest are\footnote{One could condition on a particular service $j$ using}

$$
\Delta \hat{\rho}_{kn}^k = \hat{\rho}_{kn}^k - \hat{\rho}_{1n}^k
$$

\begin{align*}
\frac{\sum_{i} \sum_{t \in T_{it}} \hat{\eta}_{it}^k b_{it}}{\sum_{i} \sum_{t \in T_{it}} b_{it}} - \frac{\sum_{i} \sum_{t \in T_{it}} \hat{\eta}_{it}^k b_{it}}{\sum_{i} \sum_{t \in T_{it}} b_{it}}
\end{align*}
as long as \( \sum_{t \in T, \tau} \hat{\eta}^k_{it} \approx 0 \) for \( \tau = 1, 3, 4 \). These \( \Delta \hat{\rho}^k_{\tau n} \) statistics and their associated t-statistics are the statistics reported in Table 8 of Section 6.2.

In Section 6.2, we also define a treated group, called a “beneficiary group,” and a control group, called a “placebo group.” Computing \( \Delta \hat{\rho}^k_{\tau n} \) separately for the beneficiary group (\( \Delta \hat{\rho}^kB_{\tau n} \)) and the placebo group (\( \Delta \hat{\rho}^kP_{\tau n} \)), we can take another difference,

\[
\Delta \hat{\rho}^{k\Delta}_{\tau n} = \Delta \hat{\rho}^kB_{\tau n} - \Delta \hat{\rho}^kP_{\tau n}.
\]

Under the null hypothesis that the treatment has no effect on outcomes, \( \Delta \hat{\rho}^{k\Delta}_{\tau n} \) should be close to zero. Again, we can simulate the distribution of \( \Delta \hat{\rho}^{k\Delta}_{\tau n} \) under \( H_0 \) and use it to construct critical values for \( \Delta \hat{\rho}^{k\Delta}_{\tau n} \). Results associated with \( \Delta \hat{\rho}^{k\Delta}_{\tau n} \) are reported in Tables 9 and 10. Finally, using methods described in the appendix, the \( \Delta \hat{\rho}^{k\Delta}_{\tau n} \) statistics are used to compare magnitudes of the effects of local paratransit characteristics on VR service effects relative to the effects of personal transportation characteristics (see the personal transportation variables in Table 2) on VR service effects. These results are reported in Tables 11 and 12.

4.2 Testing in the Presence of Measurement Error

Another significant advantage of our proposed method of exploring for transportation effects over a more traditional method of re-estimating the models in DPSS is that we can perform sensitivity analyses easily. Doing so for estimated effects is prohibitively computationally costly in this context. Such analyses indicate that our method is robust to measurement error in the imperfect measures of paratransit availability we employ (described in Section 5.3). To make the point more clearly, consider a simple linear model of the form,

\[
Y_i = W_i \beta + V_i \psi + \varepsilon_i^Y
\]

\[
\varepsilon_i^Y \sim iid (0, \sigma^2_Y),
\]

and, for simplicity, assume that \( V_i \) is a scalar and \( V_i \perp W_i \). Assume that \( V_i \) is not observed, and instead,

\[
C_i = V_i + \varphi_i
\]

is observed, \( \varphi_i \sim iid (0, \sigma^2_\varphi) \). The hypothesis of interest is

\[
H_0 : \psi = 0 \text{ vs } H_A : \psi \neq 0.
\]

Then, as is well-known, using OLS to estimate just \( \beta \) provides a consistent estimate. We can define our pseudo-LM test statistic as the correlation of residuals \( \hat{\varepsilon}_i^Y \) from the OLS regression and \( C_i \),

\[
T_G = \frac{\frac{1}{n} \sum_i \hat{V}_i \hat{\varepsilon}_i^Y}{\sqrt{\left( \frac{1}{n} \sum_i \hat{V}_i^2 \right) \left( \frac{1}{n} \sum_i \hat{\varepsilon}_i^Y \right)}}
\]
(where $\tilde{V}_i = V_i - \bar{V}$) with

\[
A = \frac{\text{plim} (T_G) = \psi A,}{\sqrt{\text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \right) \text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \psi^2 + \sigma_i^2 \right)}} > 0.
\]

Under $H_0$, $\text{plim} (T_G) = 0$, and, under $H_A$, $\text{plim} (T_G)$ and $\psi$ have the same sign.

The fact that $\text{plim} (T_G) \neq \text{Corr} (V_i, Y_i - W_i \beta)$ (because of measurement error) is not particularly important because the distribution of the test statistic $T_G$ can be simulated under the assumption of measurement error. Since $\text{plim} (T_G)$ is biased towards zero, ignoring the existence of measurement error in the distribution simulation process makes the critical value for $T_G$ higher than it should be. However, given the relatively simple structure of the effect of measurement error on $A$, it is easy to perform a sensitivity analysis of the effect of measurement error on critical values (and, therefore, power).

One also could allow for nonclassical measurement error (see, for example, Black, Berger, and Scott, 2000; Kreider and Pepper, 2007). Consider the same model as in equation (9) but assume that $V_i$ is binary and

\[
\begin{align*}
\Pr (C_i = 1 | V_i = 1) &= p \\
\Pr (C_i = 0 | V_i = 0) &= q
\end{align*}
\]

with

\[p > \frac{1}{2}, q > \frac{1}{2}\]

As before, consider the hypothesis,

\[H_0 : \psi = 0 \text{ vs } H_A : \psi \neq 0.\]

Note that $\hat{\beta}$ is consistent and

\[
\text{plim} \hat{\psi} = \psi \left( \text{plim} \frac{1}{n} \sum_i V_i \right)
\]

\[
\left[ (p + q - 1) \text{plim} \frac{1}{n} \sum_i V_i + (1 - q) \right].
\]

Alternatively, we can use OLS residuals from a regression of $Y$ on $W$ to calculate

\[
T_G = \frac{\frac{1}{n} \sum_i \tilde{C}_i \tilde{\varepsilon}_i^Y}{\sqrt{\left( \frac{1}{n} \sum_i \tilde{C}_i^2 \right) \left( \frac{1}{n} \sum_i (\tilde{\varepsilon}_i^Y)^2 \right)}}
\]

Then,

\[
\text{plim} (T_G) = R (p, q) = \sqrt{\left[ (p + q - 1) \text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \right) \right] + 1 - q \text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \right)} + \left( \frac{\sigma_i}{\psi} \right)^2 > 0. \quad (10)
\]
Figure 2 shows how $\text{plim}(T_G)$ varies with $p$ and $q$ for the arbitrary special case where $\text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \right) = 0.4$ and $\text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \right) + \left( \frac{\alpha \psi}{\psi} \right)^2 = 1.1$. The variation depends only on the numerator in equation (10) because the denominator includes neither $p$ nor $q$. The circle in Figure 2 corresponds to the case when there is no measurement error (i.e., $p = q = 1$) where

$$\text{plim} (T_G) = R(1, 1) = \sqrt{\frac{\text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \right)}{\text{plim} \left( \frac{1}{n} \sum_i \tilde{V}_i^2 \right) + \left( \frac{\alpha \psi}{\psi} \right)^2}};$$

all other points should be viewed relative to that point. Thus, if there is no correction for the measurement error and $R(p, q) > R(1, 1)$, then the critical value will be too large; if $R(p, q) < R(1, 1)$, then the critical value will be too small.

Note that performing a sensitivity analysis in this testing mode requires only adjusting critical values for different values of $p$ and $q$ in equation (10). Meanwhile, performing a similar sensitivity analysis based on estimation requires re-estimation for each combination of $p$ and $q$ of interest. The former sensitivity analysis is extremely fast, while the latter is quite expensive, especially for complex models of the sort discussed in Section 2.

Given the imperfect nature of the available paratransit data, we are concerned that measurement errors in our variables of interest may affect our results. This sensitivity analysis reveals that our method is robust to this concern for three reasons. First, as is the case in many problems, measurement error does not greatly affect our results. Figure 2 indicates that the potential bias is not large in magnitude, even as $p$ and $q$ move away from 1. Second, this point is even more salient when viewed relative to the magnitudes of the t-statistics we report in Section 6.3. The t-statistics are sufficiently large that it is unlikely that any potential bias would qualitatively affect our findings. Finally, measurement error in our data takes two different forms. Some individuals with access to paratransit may be incorrectly coded as not having paratransit...
available \((p < 1)\). Conversely, some individuals without access may be falsely reported to have access \((q < 1)\). Figure 2 indicates that the biases from these two different types of measurement error move in opposite directions. When \(p < 1\), our test statistics are biased downward; when \(q < 1\), they are biased upward. This mitigates the effects of any potential bias.

5 Data

In this section, we discuss the data used in the project. There are three major sources of data used in the project, discussed below, and then some minor ones discussed in DPSS. The first is the administrative records of all people who applied for vocational rehabilitation services for the Virginia Department of Aging and Rehabilitation Services (DARS) in 2000. The second is the administrative records of the Virginia Employment Commission (VEC) for all of the applicants to DARS for 3 years prior to DARS service until 7 years after service receipt. The VEC data provide us with quarterly information on earnings over 10 years. The last is data that we constructed on the characteristics of the various paratransit systems across Virginia over the 10 years corresponding to the VEC data. Much of the public transportation data came from Department of Transportation (US DOT) websites, but some required phone calls to providers across Virginia.

5.1 DARS Data

As already mentioned, the first source of data is the administrative records for all applicants for vocational rehabilitation services to DARS in 2000. We observe information on the vocational rehabilitation services received along with some demographic and disability/health information. We decompose the sample into three large disability groups: people with cognitive impairments (Dean et al., 2015), people with mental illness (Dean et al., 2017a), and people with physical impairments (Dean et al., 2017b).

Table 1 provides information on the proportion of people receiving services of each of the six types: diagnostic & evaluation services, training services, education services, restoration services, maintenance services, and other services. These are described in more detail in Appendix 1 and in DPSS. There is significant variation across disability groups with respect to the reported proportions. For example, training services are received by 42.6% of clients with cognitive impairments, 37.2% of people with mental illness, and 14.5% of people with physical impairments. Education services are received by only 3.4% of clients with cognitive impairments and 13.2% of people with mental illness.

Table 2 provides information about the means of selected explanatory variables for each of the three disability groups.\(^\text{17}\) There are significant differences in means for most of the variables and in ways that make sense. For example,

\(^{17}\)Complete lists of explanatory variables, along with means and standard deviations, are available in DPSS.
the proportion of people with cognitive impairments who had a special education certificate was much higher (0.332) than for people with mental illness (0.025) or people with physical impairments (0.014). Meanwhile, the means for having a driver’s license go in the opposite order.

Variables of particular interest for this study are whether the individual has a driver’s licence and whether the individual thinks that there is available transportation. Available transportation can be provided by a family member, a public transit system, or a paratransit system specializing in the transportation needs of disabled people.18 Having a driver’s license is likely the reason individuals with mental illnesses or physical impairments have access to transportation; meanwhile, the majority of people with cognitive impairments rely on alternative forms of transportation.

5.2 VEC Data

The second source of data is the administrative records from the VEC for each person in the DARS data from 3 years prior to service receipt up to 7 years after service receipt. Each quarterly observation reports quarterly earnings for the particular individual.19 Table 3 provides means for the two dependent

18See, for example, Cyra, Mulroy, and Jans (1988), Stern (1993), Franklin and Niemeier (1998), and Bearse et al. (2004) for descriptions of the transportation choice set of disabled people and how often they choose different alternatives.

19Details about these data are available in DPSS.
Table 3: Labor Market Means Across Disability Groups

<table>
<thead>
<tr>
<th></th>
<th>Cognitive Impairment</th>
<th>Mental Illness</th>
<th>Physical Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td># Obs</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Proportion Employed</td>
<td>58522</td>
<td>0.253</td>
<td>90190</td>
</tr>
<tr>
<td>Conditional log</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quarterly Earnings</td>
<td>14799</td>
<td>7.009</td>
<td>27148</td>
</tr>
</tbody>
</table>

Note: log Quarterly Earnings are conditional on being employed.

People with physical impairments have the highest employment rate (0.335) and conditional log quarterly earnings (7.692), people with mental illness are second in both dimensions, and people with cognitive impairments are third in both dimensions. The sample sizes are very large, providing precise estimates of all of the model parameters.

5.3 Transportation Data

The third type of data is information on the existence of public transportation in counties/cities across Virginia\(^{20}\) and performance data of all urban transportation systems and some rural transportation systems from the National Transit Database (NTD). For the existence information, we consulted the Transit Development Plan (TDP) of each transportation agency at Virginia DRPT (2016) to determine when the agency started providing service in each county it covers. If a transportation agency did not have a TDP available at Virginia DRPT (2016) or its TDP did not contain complete information on the history of its service, we obtained the information needed by phone call or email. We matched each public transportation provider with one or multiple cities/counties in the data set and constructed the public transportation exists variable for each city/county in each year from 1988 to 2015. Note the variable we are using is public transportation exists rather than specialized transportation exists. In Virginia, all urban areas with public transportation also have demand-response specialized transportation, and rural areas with public transportation set up their systems so that they can play both roles as required by the ADA.

One should keep in mind that availability is quite different than usage. Rosenbloom (2007) argues that, among disabled people who had available paratransit services, only 1.2% had used them at least once in the previous week. However, Bearse et al. (2003), using data from a single provider between 1984 to 1996, find that 47% of trips are taken by 7% of users, and approximately 20% of trips are taken by about 1% of users. Most of these trips are for work. Rosenbloom (2007) thinks this is a problem for paratransit systems. For us, it means that availability probably is a good measure for the potential to use transportation.

For performance data, we choose four performance data variables from US DOT (2016) relevant to our research to further analyze: per capita passen-

\(^{20}\)Virginia, unlike all other states, has independent cities that are not part of their adjacent counties.
ger miles traveled \((PMT)\), per capita vehicles operated in maximum service \((VOMS)\), per capita vehicle revenue hours \((VRH)\),\(^{21}\) and per capita unlinked passenger trips \((UPT)\).\(^{22}\) For rural transportation agencies, due to the lack of data sources available,\(^{23}\) we collected \(UPT\) data only for those agencies included in US DOT \((2016)\). Among these four types of transit performance data, \(PMT\) and \(UPT\) provide information on service consumed by passengers, and \(VOMS\) and \(VRH\) provide information on service supplied by transportation agencies.\(^{24}\) Each is a measure of a particular dimension of public transportation existence and quality. The most common measure used in the literature is \(UPT\) \((TCRP, 2013; Mattson, 2016)\). We experiment with each because we do not know a priori which might be the most relevant measures. Mattson \((2015)\) suggests that \(PMT, VRH, and UPT\) are useful measures.

Table 4 provides summary moment information on the five measures used in our analysis: \(public transportation exists, PMT, VOMS, VRH, and UPT\). The number of observations varies significantly across variables due to variation of the availability of the data in the NTD.\(^{25}\) We decompose the standard deviation of each variable into an across-county standard deviation and a within-county, across-time standard deviation. A little more than half of the variation is across-county variation. As seen in Table 5, a significant portion of the within-county standard deviation can be explained by secular growth of agencies (even relative to population). This is consistent with results in Bearse et al. \((2004)\), National Council on Disability \((2005)\), and Sapper, Goodwill, and Carapella \((2009)\) but not consistent with the cross-section results in Koffman et al. \((2007)\).\(^{26}\)

There are two shortcomings of our paratransit measures. First, there is

\(^{21}\)Revenue hours are the time when a vehicle is available to the general public and there is an expectation of carrying passengers. Vehicles operated in fare-free service are considered in revenue service. Revenue service excludes school bus service and charter service.

\(^{22}\)Passengers are counted each time they board vehicles no matter how many vehicles they use to travel from their origin to their destination and regardless of whether they pay a fare, use a pass or transfer, ride for free, or pay in some other way.

\(^{23}\)Data for rural transportation agencies are collected only if they receive funding from the federal government. The U.S. Department of Transportation started collecting information from such rural agencies only in 2007.

\(^{24}\)The data reported in US DOT \((2016)\) are not per capita. To construct per capita variables, we extract population data from US BEA \((2016)\) for each city and county in Virginia and then aggregate over cities and counties covered by each agency.

\(^{25}\)There are two issues associated with the process of collecting and cleaning data. First, some paratransit providers include multiple cities or counties in their service areas, some of which are receiving only marginal services. In such cases, if the area covered by the transportation agency is less than 50% of the area of that county, we treat that county as not covered by the agency. Second, in US DOT \((2016)\), there are several “zeros” reported. It is not clear whether any particular “zero” is a true zero or an indicator of a missing observation. In such cases, we consult the TDPs of the corresponding agencies and learn about the existence of paratransit service; hence, we treat those zero entries as indicators of missing data and exclude the associated observations from further analysis.

\(^{26}\)Koffman et al. \((2007)\) assert that it is important to include only the population that lives within 3/4 mile from fixed route systems because people living farther away are not covered according to ADA rules. This is probably a poor research choice because a) it is very difficult to measure the appropriate population and b) 48% of paratransit systems do not restrict service based on the rule \((Sapper, Goodwill, and Carapella, 2009)\).
Table 4: Public Transportation Data

<table>
<thead>
<tr>
<th>Variable</th>
<th># Obs</th>
<th>Mean</th>
<th>Std Dev Across Counties</th>
<th>Std Dev Within Counties</th>
<th>Std Dev Across Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Transportation Exists</td>
<td>133</td>
<td>0.542</td>
<td>0.398</td>
<td>0.308</td>
<td>0.503</td>
</tr>
<tr>
<td>PMT per capita</td>
<td>39</td>
<td>51.654</td>
<td>47.550</td>
<td>35.103</td>
<td>59.104</td>
</tr>
<tr>
<td>VOMS per 1K capita</td>
<td>41</td>
<td>0.266</td>
<td>0.130</td>
<td>0.088</td>
<td>0.157</td>
</tr>
<tr>
<td>VRH per capita</td>
<td>41</td>
<td>0.602</td>
<td>0.353</td>
<td>0.168</td>
<td>0.391</td>
</tr>
</tbody>
</table>

Table 5: Proportion of Within-County Variation Explained by a Time Trend

<table>
<thead>
<tr>
<th>Variable</th>
<th># Obs</th>
<th>R-Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Transportation Exists</td>
<td>3724</td>
<td>27.4%</td>
</tr>
<tr>
<td>PMT</td>
<td>3724</td>
<td>18.0%</td>
</tr>
<tr>
<td>VOMS</td>
<td>659</td>
<td>16.4%</td>
</tr>
<tr>
<td>VRH</td>
<td>657</td>
<td>33.1%</td>
</tr>
<tr>
<td>UPT</td>
<td>933</td>
<td>5.1%</td>
</tr>
</tbody>
</table>

a lack of systematic data necessary to measure the existence and quality of demand-response service in United States (Godavarthy et al., 2015). Important missing variables in the National Transportation data include “geographic coverage, days of service per week, hours of service per day, advance reservation requirements, and service eligibility” (page 3). Kittelson & Associates, Inc. et al. (2003) advise measuring the level of service for demand-response transportation services using response time, service span, reliability, on-time performance, trips not served, and travel time of demand-response transit.

Second, we have geographic information of each agency only at the county level. It is possible that public transportation does not exist in some areas of a county even though the county is recorded as having public transportation in our data set. Mattson (2016) performs a moderately-sized survey collecting data from each included agency on many details of service specific to subsets of each county. We are unable to replicate his methods or use his data for two reasons. He asks only about present agency characteristics, and we need data from 15 years earlier. Also, in the VEC and DARS data, we observe only the county of residence of each applicant. Thus, there may be occurrences of an applicant incorrectly being described as having access to public transportation in our analysis. If so, this will bias our results as described in Section 4.2.

6 Results

We start off this section with a short synopsis of the most critical results in DPSS. More discussion is available in each paper. Then, we discuss the results of the pseudo-LM tests described in Section 4.
Table 6: Employment Effects of Service Receipt by Disability

<table>
<thead>
<tr>
<th>Group</th>
<th>Cognitive Impairment</th>
<th>Mental Illness</th>
<th>Physical Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short Run Long Run</td>
<td>Short Run Long Run</td>
<td>Short Run Long Run</td>
</tr>
<tr>
<td>Diagnosis &amp; Evaluation</td>
<td>0.422 0.310</td>
<td>-0.228 -0.462</td>
<td>0.547 0.127</td>
</tr>
<tr>
<td>Training</td>
<td>0.694 0.458</td>
<td>0.631 0.541</td>
<td>0.324 0.408</td>
</tr>
<tr>
<td>Education</td>
<td>0.122 0.538</td>
<td>-0.299 -0.113</td>
<td>0.258 0.564</td>
</tr>
<tr>
<td>Restoration</td>
<td>-0.331 -0.458</td>
<td>-0.017 -0.127</td>
<td>0.515 0.357</td>
</tr>
<tr>
<td>Maintenance</td>
<td>-0.124 -0.059</td>
<td>0.054 -0.074</td>
<td>-0.007 -0.025</td>
</tr>
<tr>
<td>Other Services</td>
<td>0.383 0.419</td>
<td>0.128 0.049</td>
<td>0.385 0.269</td>
</tr>
</tbody>
</table>

6.1 Vocational Rehabilitation Service Effects

Tables 6 and 7 provide estimates of the short- and long-run effects of each DARS service on labor market outcomes, disaggregated by disability group. Table 6 shows estimates for effects on the value of being employed ($e_t$ in equation (3)), also called the employment propensity. Each number reported in the table is the difference of two estimates from the corresponding model. For example, for the cognitive impairment disability group, the long-run effect of training is 0.458; this means that $\tilde{\alpha}_{j4} - \tilde{\alpha}_{j1} = 0.458$ for the $j$ corresponding to training services. In words, the receipt of training services increases the value of being employed by 0.458 more than its value prior to receipt of services, suggesting a 18.3% increase in the probability of being employed. Analogously, for short-run effects, $\tilde{\alpha}_{j3} - \tilde{\alpha}_{j1} = 0.694$ for the $j$ corresponding to training services, suggesting a 27.8% increase in the probability of being employed. The results show that training is uniformly effective in increasing employment in the short run and long run for all three disability groups. Education services are effective for people with cognitive impairments and people with physical impairments but not for people with mental illness. Restoration is effective for people with physical impairments but not the other two groups.

Receipt of maintenance services has a negative effect on employment except for people with mental illness in the short run. This might occur because receipt of maintenance services during receipt of some other service implies a need for the maintenance service. For example, one might need child care services or transportation help to the service location. Once the service period ends, the maintenance support also ends. However, the need for the support does not end. In particular, with respect to the issues associated with this paper, to the degree that DARS pays for transportation during service receipt, it suggests that the individual might have difficulty working because she has no transportation available without the DARS transportation support.

Table 7 presents analogous results for the effect of services on short- and long-run log quarterly earnings (conditional on being employed). For example, the long-run effect of training is $\tilde{\alpha}_{j4} - \tilde{\alpha}_{j1} = 0.172$; for the $j$ corresponding to

---

27 A good rule of thumb for turning $\partial e^*/\partial y_{j\tau}$ into $\partial e/\partial y_{j\tau}$ is to multiply $\partial e^*/\partial y_{j\tau}$ by 0.4 (Amemiya, 1981).

28 This point was explained to us by Bob Schmidt.
training services; i.e., the receipt of training increases earnings, conditional on employment by 17.2%. Training and education services generally have large effects on earnings, while restoration is mixed. Maintenance services have a negative effect for people with cognitive impairments, a positive effect for people with mental illness, and mixed effects for people with physical impairments. To the degree that maintenance services imply a barrier to employment, it should be expected that they would have more of a negative effect on employment than on earnings conditional on being employed.

Our results show generally positive effects for both labor market outcomes for people with cognitive impairments and mental illness (particularly in the long run). The employment effects are generally smaller in magnitude than the earnings effects. The results for the former groups suggest that some of the negative effect of maintenance services may be due to a mobility constraint but that the effect is not more pronounced in the employment propensity than in log earnings as we would expect. Alternatively, for physically impaired people, we find a robust pattern of positive employment effects and negative wage effects. Again, the employment effects are smaller in magnitude than the earnings effects. These results are consistent with maintenance reflecting lack of viable transportation for physically impaired people. Alleviating this constraint results in an increased probability of employment and may even allow more severely disabled individuals to find employment, resulting in lower average wages. Alternatively, individuals may be willing accept reduced wage offers because their lower commuting costs offsets lower compensation.

DPSS also include two measures of transportation availability: a dummy for whether the individual has available transportation and a dummy for whether he has a driver’s license. These two variables are allowed to affect each service choice value in equation (1), employment propensity in equation (3), and log quarterly earnings in equation (5). Figure 3 displays the effects of each transportation variable on employment propensity and log quarterly earnings for each of the three disability groups. All of the effects are large and positive. The largest effects are for transportation available and has driver’s license on log quarterly earnings for cognitive impairment.\(^{29}\) All estimates are statistically significant at the 5% level. No results for service choice values are reported here.\(^{30}\)

\(^{29}\)Ermagun et al. (2016) use data that allow them to estimate the effect on public trans-
DPSS each discuss how the estimates presented here in Tables 6 and 7 translate into service-specific, disability-specific rates of return. We skip that discussion here because it is not that relevant to our major concern with transportation interactions.

6.2 Test Results for Interactions

We now turn to the results of the our primary analysis, the pseudo-LM tests. In order to determine whether our estimates of the $\Delta \rho^k_m$ statistics, defined in equation (8), indicate that the effect of public transportation is meaningful, we need a counterfactual against which to compare. For all of the tables of estimates that follow, we calculate estimates for two different groups of clients who received DARS services: the “beneficiary group” and the “placebo group.” 31 We use two alternative definitions of these groups. First, we define the beneficiary group to contain those individuals who report that they do not have access to transportation (and thus might benefit from having good transportation), and the placebo group as individuals who report the opposite (and thus are unlikely to benefit from better public transportation). 32 Second, we define people in the

---

31 While these groups are similar to traditional “treatment” and “control” groups, since we do not observe in the data whether an individual actually uses the paratransit treatment or not, we use the alternative beneficiary/placebo terminology instead to highlight the distinction. The beneficiary group contains individuals who have the potential to benefit from the intervention (due to lack of transportation). Much in the way that those who receive a placebo are hopefully unlikely to benefit from a fake treatment, those in the placebo group are unlikely to benefit from paratransit because they have alternative means of transportation.

32 Our group definitions are based on a question about whether the individual believes she has access to transportation. Individuals may or may not consider public paratransit as an
beneficiary group based on clients who received maintenance services from the VR agency, and people in the placebo group as individuals who did not receive such services. We do so after noting that DPSS frequently found the effects of maintenance services on labor market outcomes to be negative and knowing of no theoretical reason why maintenance services might actually harm clients in some way. Since transportation assistance provided by DARS is considered a maintenance service, we posit that these negative effects are the result of a correlation between this service category and an unobserved factor that DPSS do not fully control for: mobility. The beneficiary group is made up of individuals who may have required transportation assistance from the VR agency while receiving services, and the placebo group contains individuals who do not require such services because they had access to transportation. In other words, negative DPSS results for people receiving maintenance services are consistent with individuals having worse labor market outcomes once they lose access to DARS provided transportation upon completing the program. We proceed by discussing results based on each definition.

6.2.1 Interpretation of Results

For the purpose of explaining how to interpret our results, we begin by presenting a subset of our pseudo-LM test results in Table 8 before summarizing the relevant information for all specifications in Tables 9 and 10. Table 8 is based on the first definition of beneficiary and placebo groups based on reported access to transportation. The table contains estimates of the short-run changes in correlations (i.e., short-run correlation after service minus correlation prior to service) between measures of the availability and intensity of public transportation in the community (discussed in Section 5.3) and the generalized residuals associated with being employed for clients with cognitive impairments. Each row contains separate estimates of the short-run change in the correlation of the residuals with the given public transportation measure. In the first column, the estimate of $\Delta \rho_{3n}^e = 0.00$ for the beneficiary group in the first row is an estimate of a double difference of the correlation of unobserved components of employment and existence of public transportation: a) the first difference is for people who received each vocational rehabilitation services minus for people who did not; and b) the second difference is for the correlation in the short-option when they report that they “have no available transportation.” To the extent that our beneficiary group contains individuals who do not have access to paratransit, our results will be attenuated.

As in the previous case, our group definitions are imperfect due to data limitations. Individuals who receive maintenance services may or may not have actually received transportation assistance. To the extent that individuals in our beneficiary group received some other maintenance service or services besides transportation, our estimates again will be biased towards zero.

The same argument could be made about child care services. But receipt of child care services would not imply that good public transportation would help in employment outcomes.

Note that, in each specification reported, the availability measure estimates are unconditional on the intensity measures. Each intensity measure estimate is conditional on transportation being available, but it is unconditional on the other intensity measures.
run quarters after service minus the correlation in the quarters prior to service. Taken alone, the point estimate of 0.00 would suggest that the existence of public transportation has no interaction with receipt of vocational rehabilitation services on short-run labor market outcomes. But, when viewed relative to the estimate of −0.06 for people in the placebo group in the second column, we find evidence that the interaction between vocational rehabilitation service and existence of public transportation increases short-run employment probabilities for cognitively impaired clients. The third column contains the estimate of this effect of 0.06, and the fourth column contains the t-statistic associated with the difference which is statistically significant for the existence measure. The estimate used here is akin to a differences-in-differences-in-differences estimator (e.g., Gruber, 2007). Explicitly, it shows that, relative to a baseline of individuals with access to transportation, the existence of public transportation would, on average, improve the effectiveness of vocational rehabilitation services on short-run labor market outcomes for cognitively impaired individuals who do not have other means of transportation available.

### 6.2.2 Definition of Beneficiary Group: No Transportation Available

As the estimated correlations do not have a direct interpretation, the sign and magnitude of the t-statistic contains all relevant information from our estimates. Table 9 reports these t-statistics for our two outcomes (employment and conditional log quarterly earnings) for each of the three disability groups receiving treatment (people with cognitive impairments, mental illness, or physical impairments), and for both the short and long run. As in Table 8, the beneficiary group is defined as people without access to transportation. Panel A (Panel B) reports the employment (conditional log quarterly earnings) estimates of the effect of the interaction of public transportation and vocational rehabilitation service receipt. Overall, the mixed patterns of sign and significance in both panels are similar to those in Table 8.

---

**Table 8: Short-Run Employment Effects for Cognitive Impairment Disability Group**

<table>
<thead>
<tr>
<th></th>
<th>Treated Group</th>
<th>Control Group</th>
<th>Difference</th>
<th>t Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Transportation Available</td>
<td>0.00</td>
<td>−0.06</td>
<td>0.06</td>
<td>2.11**</td>
</tr>
<tr>
<td>PMT per Capita</td>
<td>−0.07</td>
<td>0.07</td>
<td>−0.15</td>
<td>−4.03##</td>
</tr>
<tr>
<td>VOMS per Capita</td>
<td>−0.05</td>
<td>0.00</td>
<td>−0.05</td>
<td>−1.48</td>
</tr>
<tr>
<td>VHR per Capita</td>
<td>−0.10</td>
<td>0.04</td>
<td>−0.14</td>
<td>−3.80##</td>
</tr>
<tr>
<td>UPT per Capita</td>
<td>−0.07</td>
<td>0.00</td>
<td>−0.07</td>
<td>−1.77#</td>
</tr>
</tbody>
</table>

Notes:
1. The treated group contains clients who do not have personal transportation available.
2. For t-statistics that are consistent with theoretical predictions, double-starred items are statistically significant at the 5% level, and single-starred items are statistically significant at the 10% level.
3. For t-statistics that are inconsistent with theoretical predictions, double-pound items are statistically significant at the 5% level, and single-pound items are statistically significant at the 10% level.
For clients with cognitive impairments (columns (1) and (2)), we find positive, significant employment effects in both the short- and long-run for the existence of public transportation but predominantly significant, negative impacts along the four intensive-margin variables (PMT, VOMS, VRH, and UPT) for employment (Panel A) and for both sets of transportation variables for conditional earnings (Panel B).

The estimates in columns (3) and (4) show that clients with a mental illness display no significant interactions between DARS service receipt and the public transportation variables in short-run employment effects, but all of the short-run conditional log quarterly earnings performance measure interactions are significant. Only UPT suggests a positive relationship in the short run, and that interaction is also positive for both employment and earnings in the long run. Other long-run interactions are predominantly negative.

The physical impairment interactions in columns (5) and (6) are mostly in line with our expectations. Results are predominantly positive and significant (save for the long-run interaction of the existence of public transportation and service receipt on conditional log earnings). PMT seems be an important determinant of the impact of public transportation on the efficacy of DARS services for people with physical impairments as effects are significant and positive for both employment and conditional log earnings and in both the short and long run.

### 6.2.3 Definition of Beneficiary Group: Received Maintenance Services

We also are interested in the relationship between DARS maintenance services and public transportation because transportation assistance provided by DARS falls in the maintenance service category; thus, it provides an alternative definition of the beneficiary and placebo groups. In particular, if a VR client

---

**Table 9: t-Statistics of the Interaction of Public Transportation Characteristics and VR Services Between Clients Who Do and Do Not Have Transportation Available, by Disability Group**

<table>
<thead>
<tr>
<th></th>
<th>Cognitive Impairments</th>
<th>Mental Illness</th>
<th>Physical Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Employment Effects</strong></td>
<td>Short-Run</td>
<td>Long-Run</td>
<td>Short-Run</td>
</tr>
<tr>
<td>Public Transportation Exists</td>
<td>2.11 **</td>
<td>2.74 **</td>
<td>0.00</td>
</tr>
<tr>
<td>PMT per Capita</td>
<td>-4.03 ##</td>
<td>-2.96 ##</td>
<td>-0.87</td>
</tr>
<tr>
<td>VOMS per Capita</td>
<td>-1.48</td>
<td>1.59</td>
<td>-0.86</td>
</tr>
<tr>
<td>VRH per Capita</td>
<td>-3.80 ##</td>
<td>-1.01</td>
<td>0.03</td>
</tr>
<tr>
<td>UPT per Capita</td>
<td>-1.77 #</td>
<td>-2.00 ##</td>
<td>1.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel B: Conditional Log Quarterly Earnings Effects</strong></th>
<th>Short-Run</th>
<th>Long-Run</th>
<th>Short-Run</th>
<th>Long-Run</th>
<th>Short-Run</th>
<th>Long-Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Transportation Exists</td>
<td>-1.64</td>
<td>-3.54 ##</td>
<td>0.10</td>
<td>-2.15 # #</td>
<td>0.83</td>
<td>-2.63 # #</td>
</tr>
<tr>
<td>PMT per Capita</td>
<td>-1.96 # #</td>
<td>-2.34 # #</td>
<td>-2.25 # #</td>
<td>-1.29</td>
<td>2.11 **</td>
<td>3.97 **</td>
</tr>
<tr>
<td>VOMS per Capita</td>
<td>-2.47 # #</td>
<td>-2.13 # #</td>
<td>-2.44 # #</td>
<td>-3.26 # #</td>
<td>1.76 *</td>
<td>0.94</td>
</tr>
<tr>
<td>VRH per Capita</td>
<td>-2.12 # #</td>
<td>-1.51</td>
<td>-2.17 # #</td>
<td>-3.27 # #</td>
<td>1.37</td>
<td>0.20</td>
</tr>
<tr>
<td>UPT per Capita</td>
<td>-1.48</td>
<td>-1.40</td>
<td>-4.13 **</td>
<td>-5.81 **</td>
<td>0.64</td>
<td>2.15 **</td>
</tr>
</tbody>
</table>

*Notes:*
1. For t-statistics whose signs are consistent with theoretical predictions, double-starred items are statistically significant at the 5% level, and single-starred items are statistically significant at the 10% level.
2. For t-statistics whose signs are inconsistent with theoretical predictions, double-pound items are statistically significant at the 5% level, and single-pound items are statistically significant at the 10% level.
was receiving transportation services from DARS during receipt of other services, this would imply that the client had no other way to travel to the service site. Thus, once receipt of other services was complete, the client would lose her transportation services and again be unable to travel, now to a potential job. In fact, negative DPSS results for people receiving maintenance services are consistent with individuals having worse labor market outcomes once they lose access to DARS-provided transportation. Table 10 is analogous to Table 9 except that we define the beneficiary and placebo groups based on whether the individual received maintenance services from DARS as opposed to whether she reported having available transportation.

The cognitive impairment results are more consistent with our expectations than in the previous specification, with 5 positive, significant interactions versus 3 negative, significant interactions in this specification (compared to 2 and 11 in the specification reported in Table 9). However, the results do not show a robust pattern. Only the $UPT$ measure has a significant interaction with VR service receipt (on employment) in both the short and long run. Similarly, only the $PMT$ measure has a significant interaction with VR service receipt on both employment and conditional log quarterly earnings (in the long run).

In a stark contrast to the results from the previous specification, 8 of the 9 significant estimates for individuals with mental illness show that public transportation has a positive interaction on labor market outcomes. Only existence of public transportation shows a negative interaction in the long run. The $VRH$ measure (in the long run) and the $UPT$ measure (in both the short and long run) are particularly robust, as effects are positive and significant for both employment and earnings.

Finally, the physical impairment results are the most consistent of any group
in either specification. All of the employment effects in Panel A indicate that public transportation increases the effect of DARS service receipt on the probability of employment in the short run, and 3 of 5 effects persist to the long run. In contrast, the conditional earnings measures in Panel B are predominantly negative, indicating that, while public transportation may increase the DARS service effect on the probability of being employed, it also decreases the DARS service effect on log quarterly earnings that DARS clients with a physical impairment can expect to receive. These negative results are consistent with paratransit making it possible for people with more significant physical limitations to participate in the labor market. Alternatively, VR clients with access to efficient paratransit may be more willing to accept low-wage jobs because their commutes are not as onerous.\footnote{There is ample evidence in the housing choice literature that individuals are willing to pay to reduce their commute times. See, for instance, Bajari and Kahn (2005, 2008), Langer and Winston (2008), and Bayer and McMillan (2012). Mayock (2016) finds evidence that this trade-off also exists in wages.}

A special supplemental survey of the NHIS focusing on transportation needs of disabled people (NHIS-D) found that, among people who had difficulties “getting around,” over 75\% reported having walking problems (a physical impairment) while only 10\% reported having cognitive impairments or mental illness (Rosenbloom, 2007). Thus, it is not so surprising that results for people with physical impairments are the strongest.

6.3 Discussion of Results

Overall, results in both tables are strongly significant, but a high proportion of results are negative, particularly in Table 9. We know of no theoretical reason why (increased) access to transportation should reduce the beneficial impacts of vocational rehabilitation service receipt on labor market outcomes. Possible empirical explanations are a) our measures of transportation systems do not adequately capture the relevant dimensions of the availability and quality of public transportation that affect the behavior of individuals in our sample; b) the measures we have are actually reflecting some type of selection and/or endogeneity effect; or c) our beneficiary/placebo group definitions are imperfect representations of people likely to benefit from paratransit and our negative estimates are spurious.

What critical dimensions might we be missing? First, we do not actually know if public transportation is available to each individual in our data; we just know it is available in the individual’s county. The public transportation system may serve a limited geographic area within the county, may operate a restrictive number of hours each day, or may have onerous reservation or proof-of-eligibility requirements that renders it impractical for individuals who otherwise might use it. This mismeasurement would bias our estimates but would not explain negative estimates.

Second, we do not know if public transportation is a feasible means of travel within the community for each individual. In other words, we do not observe
how long it takes an individual to travel via public transportation to the specific location of a potential job. More popular systems may have to coordinate more pick-ups and drop-offs, increasing travel times, decreasing reliability, and making the system more costly for individual users. This would mean that increases in our intensity measures reflect a less efficient means of transportation and could explain negative intensity results.

Section 4 included a discussion of testing in the presence of classical and nonclassical measurement error. The important source of measurement error in our results is that some people may be misclassified as to whether they have access to paratransit services. This not only causes non-classical measurement error in the binary measures of access but also in the continuous measures which depend upon access. Without any evidence to back up our belief, we guess that the magnitude of false positives and negatives are both on the order of 10%. Figure 2 suggests that the testing bias associated with that magnitude is relatively small.

With regard to selection and/or endogeneity, to result in negative estimates, it would have to be the case that the source of bias results in a) the beneficiary group’s labor market outcomes after receiving VR services being worse where there are more robust paratransit systems, b) the placebo group’s outcomes being better after receiving VR services in locations with better paratransit, or c) a combination of the two. It is not obvious why this would be the case.

A likely concern in this context is migration in response to labor market opportunities. For a negative bias due to migration to drive our theoretically inconsistent results, it would have to be the case that people without transportation are more likely to migrate in response to labor market opportunities after receiving VR training than people with transportation and that migrants move to better labor market areas that have worse paratransit systems. We know of no reason to expect people without day-to-day transportation to be more mobile in their residential choices, and better job prospects are likely to be found in urban areas, where we would expect to find better, not worse, paratransit systems.

39. While it is feasible to obtain reasonable measures of this time cost by modeling the transportation network (see Clapp, 2017), the lack of information on the home (and job) locations of individuals in our data (due to confidentiality reasons) prevents the use of this type of more detailed measure. Goodwill and Joslin (2013) update two-decade-old forecasts of paratransit demand frequently used by paratransit providers to make service decisions. They do not estimate demand; rather they use trip rates from the 2009 National Household Travel Survey for households with zero vehicles applied to a measure of the local disabled population from the U.S. Census Bureau. This is likely to be a biased measure of actual usage to the extent that nondisabled people have no vehicles.

40. Replacing the arbitrary assumptions associated with the figure with assumptions closer to the actual data do not change the statements made.

41. Schachter (2001) reports that 16.2% of moves reported in the 2000 Current Population Survey were for work-related reasons. However, the existing literature suggests that employment-driven migration may be less relevant for the population we analyze. Schachter (2001) and Amior (2017) provide evidence that low-skill workers are less likely to move for job-related reasons than their high-skill peers. Also, adults with disabilities are 6% less likely to move (1-year hazard rate = 9.3%) than people without disabilities (1-year hazard rate = 9.9%) (Mateyka, 2015).
Alternatively, there is evidence that employers favor low-wage employees with shorter commutes (Phillips, 2016). For this type of endogeneity to confound our double-difference methodology, it would have to be the case that the labor market outcomes of people with disabilities are bounded from below (e.g., unemployment or minimum wage) prior to receipt of VR services. If this is the case, both beneficiary and placebo groups may appear to start out on equal footing when, in reality, people without transportation have an additional hurdle to overcome. This means that people in the placebo group who have personal transportation receive a greater boost from services from DARS than people in the beneficiary group with longer paratransit commutes. For this situation to result in negative estimates, it would have to be the case that all of the following are true: a) prior to receipt of services, both groups face poor job prospects because of their disabilities; b) people in the beneficiary group also face poor prospects due to their limited mobility; c) this additional limitation does not disadvantage the beneficiary group further; and d) after receiving VR services, the disability penalty for both groups decreases, but, for the beneficiary group, this improvement is offset by the now binding mobility penalty. For such a situation to result in negative estimates, it would also have to be the case that the true effect of paratransit on labor market outcomes is close to zero so that the interaction results are primarily driven by the difference in labor market outcomes between the beneficiary and placebo groups in the post period. Although possible, it seems unlikely that all of these conditions would hold and drive our results.

We view the final listed explanation, imprecise beneficiary/placebo group definitions, as a likely contributor to some of our unintuitive empirical results. Ultimately, while there are negative, significant estimates in both Tables 9 and 10, we find the set of results based on the transportation available beneficiary/placebo group definitions (Table 9) to be implausible but the overall set of estimates based on the maintenance services group definitions in Table 10 to be reasonable. Using the latter definitions, only the earnings interactions for individuals with physical impairments are consistently negative and significant, but the employment effects for those individuals are almost all positive and significant. The pattern of results across the two definitions suggests that people with transportation available are not a good placebo group for people without transportation available. However, the people who received maintenance services were a relatively similar group of individuals to people who did not prior to the receipt of VR services.

6.4 Economic Significance

Therefore, we proceed by comparing interactions based on the maintenance services receipt definition to the estimates from DPSS to provide a sense of economic significance relative to other interventions. Tables 11 and 12 contain transportation variable effect estimates from both DPSS and our study (denoted as either DPSS or CSY in the first column of the tables). Table 11 reports the effect of a change in the given transportation measure on employment propen-
Transportation available
Has driver’s license
Public transportation exists
PMT per capita
VOMS per capita
VRH per capita
UPT per capita

<table>
<thead>
<tr>
<th>Note:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) DPSS effects of transportation variables do not distinguish between the short and long run.</td>
</tr>
<tr>
<td>2) Both DPSS and CSY effects are calculated by scaling to a 1-standard deviation change in variable in the estimated double-difference correlations to make them comparable to the DPSS effects. See the appendix for details.</td>
</tr>
</tbody>
</table>

In order to provide a better sense of the magnitudes of our estimates relative to those related to the personal transportation variables in DPSS, we also plot the DPSS effects alongside the long-term CSY effects from Tables 11 and 12 in Figures 4 and 5. The figures also plot the 95% confidence interval for each effect to highlight which effects are significant.

For individuals with cognitive impairments, the “Long Run” column in Table 11 and the first grouping in Figure 4 show that two out of the five paratransit estimates (PMT and UPT) indicate significant, positive employment effects that are of similar magnitudes to the transportation measures from DPSS (reporting that one had access to transportation or a driver’s license). Additionally, the
Table 12: log Quarterly Earnings Effects of Transportation Variables by Disability

<table>
<thead>
<tr>
<th>Disability Group</th>
<th>Short Run</th>
<th>Long Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Impairment</td>
<td>DPSS</td>
<td>Has Driver's License</td>
</tr>
<tr>
<td></td>
<td>0.305**</td>
<td>0.141 **</td>
</tr>
<tr>
<td>Mental Illness</td>
<td>0.039**</td>
<td>0.132 **</td>
</tr>
<tr>
<td>Physical Impairment</td>
<td>0.070* *</td>
<td>0.070 **</td>
</tr>
</tbody>
</table>

Notes:
1) DPSS effects of transportation variables do not distinguish between the short and long run.
2) Both DPSS and CSY effects are calculated by scaling to a 1-standard deviation change in variable in the estimated double-difference correlations to make them comparable to the DPSS effects. See the appendix for details.
“Long Run” column in Table 12 and the first grouping in Figure 5 indicate that the long-run effects of the PMT and VOMS paratransit measures on the change in conditional log quarterly earnings after receipt of VR services are both positive, significant, and of the same order of magnitude as the DPSS measures. The patterns suggest that, for people with cognitive impairment, paratransit is a viable alternative to personal transportation. Improved paratransit leads to statistically and economically significant increases in both employment propensities and conditional earnings.

The second groupings in Figures 4 and 5 present analogous results for clients with mental illness. Figure 4 shows that the personal transportation effects are larger than the paratransit effects which are not statistically significant or negative. However, the conditional log quarterly earnings effects for these measures in Figure 5 are of similar or greater magnitudes than the DPSS effects. The long-run effect of an increase in UPT is roughly 2.5 times larger in magnitude than the has driver’s license effect and 10 times larger than the transportation available effect. These results suggest that, for people with mental illness, improved paratransit systems may improve long-run earnings outcomes. This increase in earnings (but not employment) is consistent with a more flexible, more reliable system allowing a VR client to work longer hours and/or pick up additional shifts. These work hours effects may also compound if they translate into increased productivity and higher wages. It is not fully clear why paratransit does not also have an effect on the employment propensities of this group of individuals. One possibility is that clients with mental illness receive more non-transportation related maintenance services than their peers in the other disability groups. If so, our beneficiary group contains more individuals who do not benefit from paratransit than for the other groups and our mental illness results are biased. For this to be the case given that we find strong earnings effects, the true earnings effects would have to be small in magnitude, while the true conditional earnings effects would have to be much larger. This is because the beneficiary/placebo definitions are the same for both the employment and earnings groups, so both sets of estimates would be biased in this scenario.

Finally, the third grouping in each figure displays the relative magnitudes of the employment and conditional log earnings results for people with physical impairments. Three of the five earnings effects (PMT, VHR, and UPT) are statistically significant and positive. All three of these employment effects have magnitudes that are bracketed by the corresponding DPSS effects. Table 11 shows that people with physical impairments receive nontrivial, short-run employment benefits from improved paratransit, and there is some evidence are two such employment effects across all of our measures. These effects are a departure from the overall pattern of employment effects, and we are not sure what would result in these unintuitive results given that pattern.

Although the 95% confidence interval for the PMT effect in Figure 5 contains zero, Table 12 indicates that the 90% confidence interval does not.

Table 1 provides evidence that is consistent with this possibility but does not provide conclusive evidence. A greater percentage (30.1%) of clients with a mental illness receive maintenance services compared to those with a cognitive or physical impairment (25.1% and 15.1%, respectively).
that those effects persist to the long run. Figure 5 shows that those gains are offset by robust, negative conditional earnings effects in the long run that are similar in magnitude to the personal transportation effects. The negative log earnings effects could occur because paratransit increases the likelihood that people with more severe physical impairments are able to obtain jobs after the receipt of DARS services (selection effect). Alternatively, paratransit may make VR clients more willing to accept lower-paying jobs because their commuting costs are reduced (compensating differential effect).

Taken together, the magnitude of the estimates in the tables and figures provides evidence that paratransit is a potentially important determinant of the labor market outcomes of VR clients relative to the effects of other forms of transportation. More precisely, the pattern of estimated existence and efficacy effects suggests that just having paratransit in one’s community is not helpful but having good paratransit in one’s community is.\textsuperscript{45} None of the public transportation exists effects are positive and statistically significant, but at least two system efficacy measures have a positive, significant effect on employment and/or earnings for each of our three disability groups. In other words, we find that efficient, high-quality paratransit has economically and statistically significant effects on the employment outcomes of individuals cognitive impairment, mental illness, and physical impairment. That these effects are not trivial in magnitude relative to those associated with measures of personal transportation is a striking result, in and of itself. Traveling by automobile is more flexible than traveling by public transportation, and Phillips (2016) shows that employers take the characteristics of employee commutes into account when making personnel decisions. Additionally, use of personal transportation may be associated with less-severe limitations among people in the VR population. Both arguments suggest that the personal transportation effects should dominate their paratransit counterparts, but this is not the case.

The lack of robust patterns and the existence of significant, negative effects in some of our specifications suggest that better measures of the availability, efficacy, and use of paratransit are needed to measure the role that paratransit plays in the labor market outcomes of disabled people. At the same time, our analysis implies that further study of this topic is warranted and has the potential to have important policy impacts. Our findings suggest that VR programs may be able to improve the labor market outcomes of their clients and/or reduce program expenditures by helping clients make better use of existing paratransit services in their communities.

\textsuperscript{45}We define “good,” or high-quality, paratransit by defining its converse. When paratransit systems do not function well, riders complain of late and missing vehicles, slow service, rude dispatchers, long waits to schedule trips, lack of responsiveness to complaints, inexperienced drivers, and inappropriate behavior by drivers (National Council on Disability, 2005). All of these issues can render a system unusable for VR clients who would otherwise benefit from paratransit services.
7 Conclusions

Results from previous work by DPSS suggest that maintenance services (such as transportation assistance) provided by VR agencies have negative effects on labor market outcomes. These findings suggest that VR agencies face a problem. They provide numerous services, at significant expense, for many clients who have access to transportation only through the VR agency. Due to this constraint, such VR clients have limited opportunity to achieve the ultimate goal of their VR training: finding gainful, long-term employment. In theory, paratransit systems should ameliorate transportation problems for disabled people. But there is significant variation in the availability and efficacy of paratransit systems (e.g., Mattson, 2015; Denson, 2015), so not all VR clients have access to paratransit that is of sufficient quality to be of use (if they have access to paratransit services at all).

In this work, we assess the impact that access to and the efficacy of public paratransit has on VR program effectiveness. Our results suggest that people with physical impairments who need transportation assistance while receiving VR services benefit from the availability of high-quality paratransit services after VR services end. The employment propensities for that group increase, but their conditional earnings also decline on average. The positive employment effect can be interpreted in a straightforward manner as beneficial to VR clients, but earnings effects are much more nuanced because they are conditional on employment. Viewed together, this combination of results is consistent with high-quality paratransit allowing people with more significant impairments to find employment. Alternatively, high-quality paratransit may act like a compensating differential. Individuals deciding whether they are willing to accept a job offer will do so only if the benefits (primarily, earnings) exceed the costs of employment (including the commuting costs). If access to better paratransit decreases the cost of employment, then more individuals will be willing to accept a given offer, and those on the margin will be willing to work for lower wages. These employment effects are extremely strong in the short run, though many also persist to the long run.

There is evidence that high-quality paratransit improves the labor market outcomes of people with a cognitive impairment or mental illness, but this evidence is less robust than for those with physical impairments. Our results for these VR clients display patterns that suggest paratransit affects each group differently. Why paratransit effects and the mechanisms they suggest differ by disability group is left as an open question. Exploration of these differences with more detailed data is an exciting avenue for future research. For people with a cognitive impairment, estimated effects are similar in magnitude to personal transportation effects and often statistically significant. This suggests that paratransit is an attractive option for most individuals with this condition. For individuals with mental illness, our results show that quality paratransit increases conditional earnings but not employment propensities. We do not know why paratransit has no effect on employment propensities for this group of VR clients. Again, interpreting the earnings effects conditional on the employment
results, this combination of effects is consistent with better transportation allowing individuals with these conditions to work more hours and/or earn higher wages.

Taken together, our results beg for continued study of this topic; it is likely to be informative to both VR and paratransit policymakers. We offer several suggestions for future work. First, better collected data on the availability and quality of paratransit is necessary. Such data likely would improve the accuracy of our estimates, provide greater insight into which elements of effective paratransit systems (reliability, flexibility, speed, etc.) improve employment outcomes, and understand the mechanisms through which those characteristics operate. An unresolved issue is how to measure the relevant paratransit service area, but data on the origins and destinations of paratransit trips for numerous agencies (Deka and Gonzales, 2014) could inform this topic.

Second, ideal data for this research question would link VR agency and employment information (as in DPSS) to information on available transportation options for clients. One way to obtain such data would be to have VR agencies collect this transportation information from their clients. Alternatively, researchers could geocode the home and work locations of VR clients (subject to appropriate confidentiality protocols). Researchers could then use this geographic information to determine the public commuting options available and their associated costs (Clapp, 2017).

Finally, while our work provides evidence of a new margin for VR agencies to explore, it does not conclusively show that relaxing transportation constraints is a cost effective way for agencies to improve the labor market outcomes of their clients. To address this question, future work will need to conduct a full cost/benefit analysis of the provision of VR services and paratransit. Doing so will require coupling estimates of the costs of providing both VR and paratransit services with work (similar to this project) that translates both VR and paratransit services into employment and wage effects.

Overall, we interpret our findings as evidence that access to high-quality paratransit plays an important role in determining the labor market outcomes of VR clients. This suggests several possible policy improvements for providers of both VR and paratransit services. First, the regularity of commutes suggests that there may be ways for paratransit providers to share costs with other, non-disabled potential riders along the route. Second, it may be worthwhile for VR agencies to continue to provide maintenance transportation in the short run, giving their clients time after finding a job to secure a more permanent means of transportation. This is especially important since our results indicate that transportation availability is crucial in the short run. Alternatively, VR agencies may wish to explore creative solutions to their clients’ transportation constraints. These might include adding driver’s education to the menu of available VR services, facilitating the formation of carpool arrangements between clients (both current and former), or developing relationships with ride-sharing companies that would be willing to transport clients to work (or mass-transit hubs) pro bono or at a reduced rate.
8 Appendices

Appendix 1: Vocational Rehabilitation Services

The definitions in this appendix are taken, word-for-word, from Dean et al. (2015).

- **Diagnosis & evaluation** are provided at intake in assessing eligibility and developing an IPE.
- **Training** includes vocationally-oriented expenditures for on-the-job training, job coach training, work adjustment, and supported employment.
- **Education** includes tuition and fees for a GED (graduate equivalency degree) program, a vocational or business school, a community college, or a university.
- **Restoration** covers a wide variety of medical expenditures including dental services, hearing/speech services, eyeglasses and contact lenses, drug and alcohol treatments, psychological services, surgical procedures, hospitalization, prosthetic devices, and other assistive devices.
- **Maintenance** includes cash payments to facilitate everyday living and covers such items as transportation, clothing, motor vehicle and/or home modifications, and services to family members.
- **Other services** consists of payments outside of the previous categories such as for tools and equipment.

Appendix 2: Methodology to Construct Correlation Coefficients across Data Sets

Consider a latent variable model,\(^ {16}\)

\[
\begin{align*}
\xi^*_i &= W_i \delta + \zeta_i, \quad i = 1, 2, \ldots, n, \\
\zeta_i &\sim iidN(0, 1), \\
\xi_i &= \Upsilon (\xi^*_i)
\end{align*}
\]

where \(\xi_i\) is observed, \(\xi^*_i\) is unobserved, and \(\Upsilon (\xi^*_i)\) is a (possibly) nonlinear function of \(\xi^*_i\). Equations (3) and (4) provide a binary choice example where \(\Upsilon (\xi^*_i) = 1 \quad (\xi^*_i > 0)\).\(^ {47}\) Equation (5) provides a linear model example where,

\(^{16}\)The normality assumption is not necessary here and is made only for concreteness.

\(^{47}\)Equations (1) and (2) also provide such an example, but they are not relevant to the exercise described in this appendix.
implicitly, \( \Gamma (\xi_i^*) = \xi_i^* \) and the restriction on the \( \text{Var} (\xi_i) = 1 \) can be relaxed. Define the generalized residual (Gourieroux et al., 1987) for the model as
\[
\hat{\zeta}_i = E (\xi_i^* - W_i \delta \mid \xi_i) .
\]

Next, consider some other random variable
\[
P_i \sim iid (\mu_p, \sigma_p^2)
\]
with
\[
E \left( \zeta_i \mid P_i \right) = \rho_{\zeta_i P} (P_i - \mu_p), \tag{11}
\]
(at least locally, i.e., when \( P_i - \mu_p \) is small. In the case of this paper, \( P_i \) should be thought of as one of the paratransit variables included in equations (6) and (7) and described in Section 5.3. Let
\[
\hat{\rho}_{\zeta_i P} = \frac{n^{-1} \sum_i \zeta_i (P_i - \mu_p)}{n^{-1} \sum_i (P_i - \mu_p)^2}
\]
be a consistent estimator of \( \rho_{\zeta_i P} \). Then, by definition, the estimated correlation of the generalized residuals and \( P \) is
\[
\text{Corr} (\zeta_i, P) = \frac{n^{-1} \sum_i \zeta_i (P_i - \mu_p)}{\sqrt{\left(n^{-1} \sum_i \zeta_i^2 \right) \left(n^{-1} \sum_i (P_i - \mu_p)^2 \right)}}.
\]
Using equation (12), we can write
\[
\hat{\rho}_{\zeta_i P} = \frac{n^{-1} \sum_i \zeta_i (P_i - \mu_p)}{\sqrt{\left(n^{-1} \sum_i (P_i - \mu_p)^2 \right) \left(n^{-1} \sum_i \zeta_i^2 \right)}} \frac{\sqrt{n^{-1} \sum_i \zeta_i^2}}{\sqrt{n^{-1} \sum_i (P_i - \mu_p)^2}}
\]
\[
= \text{Corr} \left( \zeta_i, P \right) \frac{\sigma_{\zeta}}{\sigma_p}.
\]
Then, using estimates of \( \text{Corr} \left( \zeta_i, P \right) \), \( \sigma_{\zeta} \), and \( \sigma_p \), we can evaluate \( E \left( \zeta_i \mid P_i \right) \) from equation (11) and get an estimate of the marginal effect of \( P_i \) on \( \xi_i^* \). One should recognize that this is only a first order approximation of the effect except in restricted cases where equation (11) holds no matter the size of \( P_i - \mu_p \), e.g., \( \left( \zeta_i, P_i \right) \sim iidN \left( 0, \Psi \right) \) with covariance matrix \( \Psi \).

\[\footnote{This occurs globally for a large class of models where}
\[
\begin{pmatrix}
\hat{\zeta}_i \\
P_i
\end{pmatrix}
= \begin{pmatrix}
0 \\
\mu_p
\end{pmatrix} \mp \begin{pmatrix}
\nu_1 \\
\nu_2
\end{pmatrix}
\]
\[\text{where } \nu = (\nu_1, \nu_2)' \text{ has a zero mean and a finite covariance matrix.}\]
References


[34] Koffman, D., R. Weiner, and D. Raphael (2003). *The Impact of Federal Programs on Transportation for Older Adults*. AARP Public Policy Institute, Washington, DC.


[40] Mattson, Jeremy (2016). *Estimating Ridership of Rural Demand-Response Transit Services for the General Public*. United States, Department of Transportation Report 21177060-NCTR-NDSU08, Small Urban and Rural Transit Center, Upper Great Plains Transportation Institute, North Dakota State University, Fargo, ND.


