Interactions of Public Paratransit and Vocational Rehabilitation

Christopher M. Clapp
University of Chicago

Steven Stern
Stony Brook University

Dan Yu*
University of Minnesota

April 2019

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, 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).[1] 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.[2]

*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.

[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).

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).[3] Additionally, the availability of these systems varies widely: the national proportion of counties providing demand-response service in 2013 was 79% (Mattson, 2015).[4] 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.[5] If so, the availability and 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, 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.

[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 Disability, 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.

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 for individuals with two of the three different categories of disabilities analyzed in DPSS. We separately test for the effectiveness of paratransit on employment and conditional log quarterly earnings in both the short and long run. For VR clients with cognitive impairments, 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 find at best suggestive evidence of positive effects. Our most robust findings are for individuals with physical impairments. For those VR clients, quality paratransit is associated with significant 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 otherwise 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 VR and transportation programs in improving outcomes for people with disabilities. Rosenbloom (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 programs (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 Gonzales, 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

[7] Of course, one should be aware of possible post-hoc testing bias.

for the Federal Transit Administration.[8] Rather than basing their evaluations of these programs on the impact they had on actual outcomes relevant to the targeted population, CT provides 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 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.[10] Let y^*_{ij} be the (latent) value to client i of receiving service j, $j = 1, 2, .., J$. There are $J = 6$ available services.[11] We assume that

$$y^*_{ij} = X_i\gamma^y_j + u^y_{ij} + \varepsilon^y_{ij} \tag{1}$$

where X_i is a vector of explanatory variables described in Section 5.1, $u^y_i = (u^y_{i1}, u^y_{i2}, .., u^y_{iJ})'$ is a vector of unobserved heterogeneity errors with a complex joint distribution function described in DPSS, and $\varepsilon^y_{ij} \sim$

[8] The 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.

[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.

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

[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.

$iidG_y(\cdot)$ is a person/choice-specific (idiosyncratic) error. Let

$$y_{ij} = 1\left(y_{ij}^* > 0\right) \tag{2}$$

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

$$e_{it}^* = Z_{it}\gamma^e + \sum_{j=1}^{J} y_{ij} \sum_{\tau=1}^{4} 1\left(t \in T_{\tau i}\right)\alpha_{j\tau}^e + u_{it}^e + \varepsilon_{it}^e \tag{3}$$

be the (latent) value of working where Z_{it} is a vector of explanatory variables similar to X_i but possibly time-varying, u_i^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.[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 $\left(\alpha_{j3}^e - \alpha_{j1}^e\right)$ and long-run effect $\left(\alpha_{j4}^e - \alpha_{j1}^e\right)$ of each service j. Let

$$e_{it} = 1\left(e_{it}^* > 0\right) \tag{4}$$

be an indicator for i working in period t.

Finally, let

$$w_{it} = Z_{it}\gamma^w + \sum_{j=1}^{J} y_{ij} \sum_{\tau=1}^{4} 1\left(t \in T_{\tau i}\right)\alpha_{j\tau}^w + u_{it}^w + \varepsilon_{it}^w \tag{5}$$

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

$$\begin{aligned}
Q_{ij}^y\left(u_{ij}^y, \theta\right) &= -X_i\gamma_j^y - u_{ij}^y, \\
Q_{ij}^e\left(u_{ij}^e, \theta\right) &= -Z_{it}\gamma^e - \sum_{j=1}^{J} y_{ij} \sum_{\tau=1}^{4} 1\left(t \in T_{\tau i}\right)\alpha_{j\tau}^e - u_{it}^e, \\
Q_{it}^w\left(u_{it}^w, \theta\right) &= w_{it} - Z_{it}\gamma^w - \sum_{j=1}^{J} y_{ij} \sum_{\tau=1}^{4} 1\left(t \in T_{\tau i}\right)\alpha_{j\tau}^{ew} - u_{it}^w.
\end{aligned}$$

The likelihood contribution for observation i is

$$L_i(\theta) = \int \left[\prod_{j=1}^{J} L_{ij}^y\left(u_{ij}^y, \theta\right)\right]\left[\prod_t L_{it}^{ew}\left(u_{it}^e, u_{it}^w, \theta\right)\right] dF\left(u_i, \theta\right)$$

[12] We assume away multiple periods of service.

where θ is the vector of parameters to estimate, $u_i \sim iidF(\cdot,\theta)$ is the vector of errors for i,

$$L_{ij}^{y}\left(u_{ij}^{y},\theta\right)=\left[1-G_y\left(Q_{ij}^{y}\left(u_{ij}^{y},\theta\right)\right)\right]^{y_{ij}}G_y\left(Q_{ij}^{y}\left(u_{ij}^{y},\theta\right)\right)^{1-y_{ij}}$$

is the likelihood component for receipt (or not) of service j, and

$$L_{it}^{ew}\left(u_{it}^{e},u_{it}^{w},\theta\right)=\begin{cases} G_e\left(Q_{it}^{e}\left(u_{it}^{e},\theta\right)\right) & \text{if } e_{it}=0\\ \left[1-G_e\left(Q_{it}^{e}\left(u_{it}^{e},\theta\right)\right)\right]g_w\left(Q_{it}^{w}\left(u_{it}^{w},\theta\right)\right) & \text{if } e_{it}=1\end{cases}$$

is the likelihood component for labor market outcomes, employment and quarterly earnings. The simulator of $L_i(\theta)$ is

$$\widetilde{L}_i(\theta)=\frac{1}{R}\sum_{r=1}^{R}\left[\prod_{j=1}^{J}L_{ij}^{y}\left(u_{ij}^{yr},\theta\right)\right]\left[\prod_{t}L_{it}^{ew}\left(u_{it}^{er},u_{it}^{wr},\theta\right)\right]$$

where $u_i^r=\left(\left\{u_{ij}^{yr}\right\}_{j=1}^{J},\left\{u_{it}^{er},u_{it}^{wr}\right\}_t\right)$ 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,

$$e_{it}^{*}=Z_{it}\gamma^{e}+\sum_{j=1}^{J}y_{ij}\sum_{\tau=1}^{4}1\left(t\in T_{\tau i}\right)\left(\alpha_{j\tau}^{e}+\lambda_{j\tau}^{e}b_{it}\right)+u_{it}^{e}+\varepsilon_{it}^{e}, \tag{6}$$

$$w_{it}=Z_{it}\gamma^{w}+\sum_{j=1}^{J}y_{ij}\sum_{\tau=1}^{4}1\left(t\in T_{\tau i}\right)\left(\alpha_{j\tau}^{w}+\lambda_{j\tau}^{w}b_{it}\right)+u_{it}^{w}+\varepsilon_{it}^{w} \tag{7}$$

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

$$\lambda_j=\left\{\lambda_{j\tau}^{e},\lambda_{j\tau}^{w}\right\}_{\tau=1}^{4},$$

λ_{js} as a subset of λ_j of some interest, and $\lambda_{jc}=\lambda_j\backslash\lambda_{js}$ as the complement of λ_{js}. We want to test

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

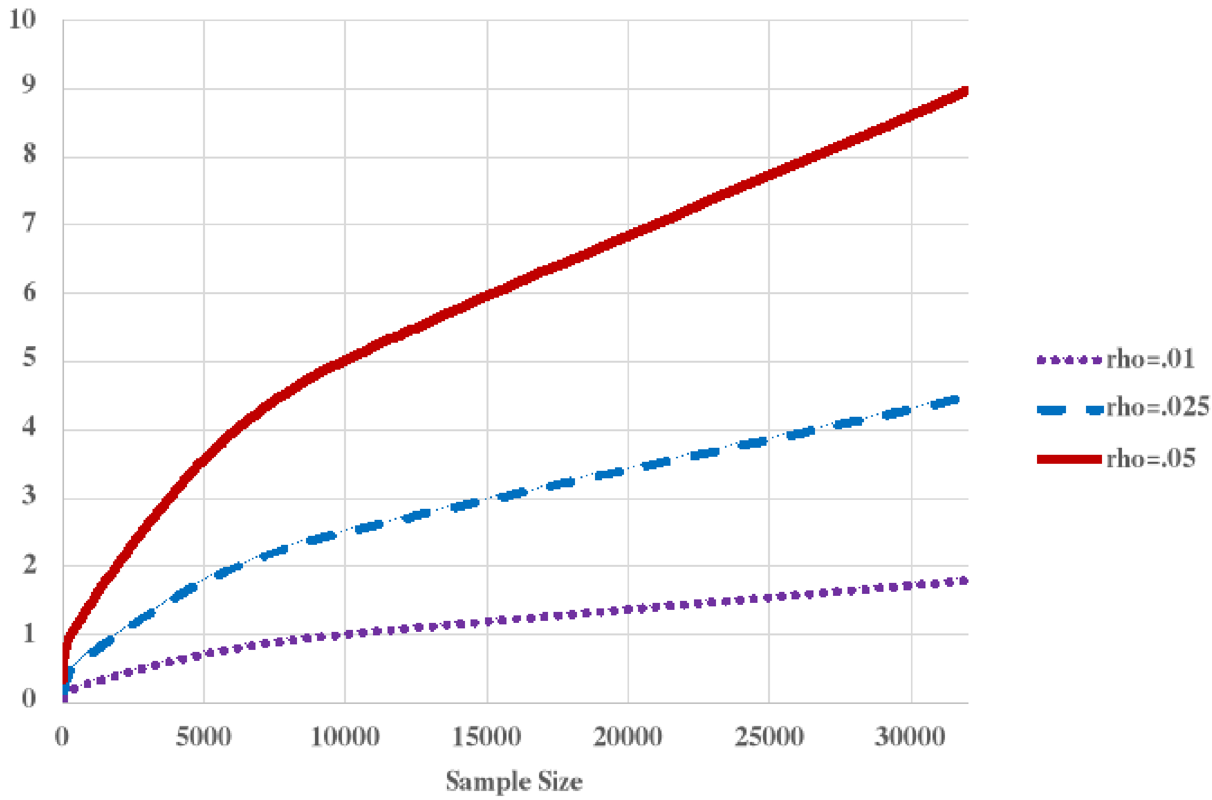
For some problems, we can define $\lambda=\{\lambda_j\}_{j=1}^{J}$ and decompose λ into λ_s and $\lambda_c=\lambda\backslash\lambda_s$. Note that, under H_0 and appropriate specification of $(\lambda_{js},\lambda_{jc})$, equations (6) and (7) simplify to equations (3) and (5). Using a Lagrange Multiplier (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 $\partial e_{it}^{*}/\partial\lambda_{js}$ and $\partial w_{it}/\partial\lambda_{js}$ (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 outcomes not "explained" by the included covariates.[14] Let $\widehat{\rho}_{\tau n}^{k}$ be the

[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.

[14] In this case, one might want to focus on just positive derivatives.

Figure 1: Test Statistic



sample correlation for $k = e, w$, constructed with n observations.[15] Then, under H_0,

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

for large n. Figure 1 shows how the test statistic varies for three relatively small values of $\widehat{\rho}_{\tau n}^{k}$ as n gets large. Note that, for employment, $n \approx 32000$, and, for log earnings, $n \approx 10000$. Checkovich 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 $\widehat{\eta}_{it}^{k}$ as the generalized residual of type $k = e, w$ for i in period t. Then the test statistics of interest are[16]

$$\frac{\sum_i \sum_{t \in T_{\tau i}} \widehat{\eta}_{it}^{k} b_{it}}{\sum_i \sum_{t \in T_{\tau i}} b_{it}} - \frac{\sum_i \sum_{t \in T_{1i}} \widehat{\eta}_{it}^{k} b_{it}}{\sum_i \sum_{t \in T_{1i}} b_{it}}$$

for $\tau = 3, 4$ which is equivalent to

$$\Delta \widehat{\rho}_{\tau n}^{k} = \widehat{\rho}_{\tau n}^{k} - \widehat{\rho}_{1n}^{k} \tag{8}$$

[15] 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.

[16] One could condition on a particular service j using

$$\frac{\sum_i y_{ij} \sum_{t \in T_{\tau i}} \widehat{\eta}_{it}^{k} b_{it}}{\sum_i y_{ij} \sum_{t \in T_{\tau i}} b_{it}} - \frac{\sum_i y_{ij} \sum_{t \in T_{1i}} \widehat{\eta}_{it}^{k} b_{it}}{\sum_i y_{ij} \sum_{t \in T_{1i}} b_{it}}.$$

as long as $\sum_{t\in T_\tau} \widehat{\eta}_{it}^k \approx 0$ for $\tau = 1,3,4$. These $\Delta\widehat{\rho}_{\tau n}^k$ 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 $\Delta\widehat{\rho}_{\tau n}^k$ separately for the beneficiary group ($\Delta\widehat{\rho}_{\tau n}^{kB}$) and the placebo group ($\Delta\widehat{\rho}_{\tau n}^{kP}$), we can take another difference,

$$\Delta\widehat{\rho}_{\tau n}^{k\Delta} = \Delta\widehat{\rho}_{\tau n}^{kB} - \Delta\widehat{\rho}_{\tau n}^{kP}.$$

Under the null hypothesis that the treatment has no effect on outcomes, $\Delta\widehat{\rho}_{\tau n}^{k\Delta}$ should be close to zero. Again, we can simulate the distribution of $\Delta\widehat{\rho}_{\tau n}^{k\Delta}$ under H_0 and use it to construct critical values for $\Delta\widehat{\rho}_{\tau n}^{k\Delta}$. Results associated with $\Delta\widehat{\rho}_{\tau n}^{k\Delta}$ are reported in Tables 9 and 10. Finally, using methods described in the appendix, the $\Delta\widehat{\rho}_{\tau n}^{k\Delta}$ 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,

$$\begin{aligned} Y_i &= W_i\beta + V_i\psi + \varepsilon_i^Y \\ \varepsilon_i^Y &\sim iid\left(0, \sigma_Y^2\right), \end{aligned} \tag{9}$$

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, $\varphi_i \sim iid\left(0, \sigma_\varphi^2\right)$. 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 β provides a consistent estimate. We can define our pseudo-LM test statistic as the correlation of residuals $\widehat{\varepsilon}_i^Y$ from the OLS regression and C_i,

$$T_G = \frac{\frac{1}{n}\sum_i \widetilde{V}_i \widehat{\varepsilon}_i^Y}{\sqrt{\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right)\left(\frac{1}{n}\sum_i \widehat{\varepsilon}_i^Y\right)}}$$

(where $\widetilde{V}_i = V_i - \overline{V}$) with

$$plim\left(T_G\right) = \psi A,$$

$$A = \frac{plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right)}{\sqrt{plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2 + \sigma_\varphi^2\right) plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\psi^2 + \sigma_Y^2\right)}} > 0.$$

Under H_0, $plim\left(T_G\right) = 0$, and, under H_A, $plim\left(T_G\right)$ and ψ have the same sign. The fact that

$$plim\left(T_G\right) \neq Corr\left(V_i, Y_i - W_i\beta\right)$$

(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 $plim\left(T_G\right)$ 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

$$\begin{aligned}\Pr\left(C_i = 1 \mid V_i = 1\right) &= p \\ \Pr\left(C_i = 0 \mid V_i = 0\right) &= q\end{aligned}$$

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 $\widehat{\beta}$ is consistent and

$$plim\widehat{\psi} = \psi \frac{plim\frac{1}{n}\sum_i V_i}{\left[(p+q-1)\, plim\frac{1}{n}\sum_i V_i\right] + (1-q)}.$$

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

$$T_G = \frac{\frac{1}{n}\sum_i \widetilde{C}_i \widehat{\varepsilon}_i^Y}{\sqrt{\left(\frac{1}{n}\sum_i \widetilde{C}_i^2\right)\left(\frac{1}{n}\sum_i \left(\widehat{\varepsilon}_i^Y\right)^2\right)}}$$

Then,

$$plim\left(T_G\right) = R\left(p,q\right) = \sqrt{\frac{\left[(p+q-1)\, plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right)\right] + 1 - q}{plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right) + \left(\frac{\sigma_Y}{\psi}\right)^2}} > 0. \tag{10}$$

Figure 2 shows how $plim\left(T_G\right)$ varies with p and q for the arbitrary special case where $plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right) = 0.4$ and $plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right) + \left(\frac{\sigma_Y}{\psi}\right)^2 = 1.1$. 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 (ie, $p = q = 1$) where

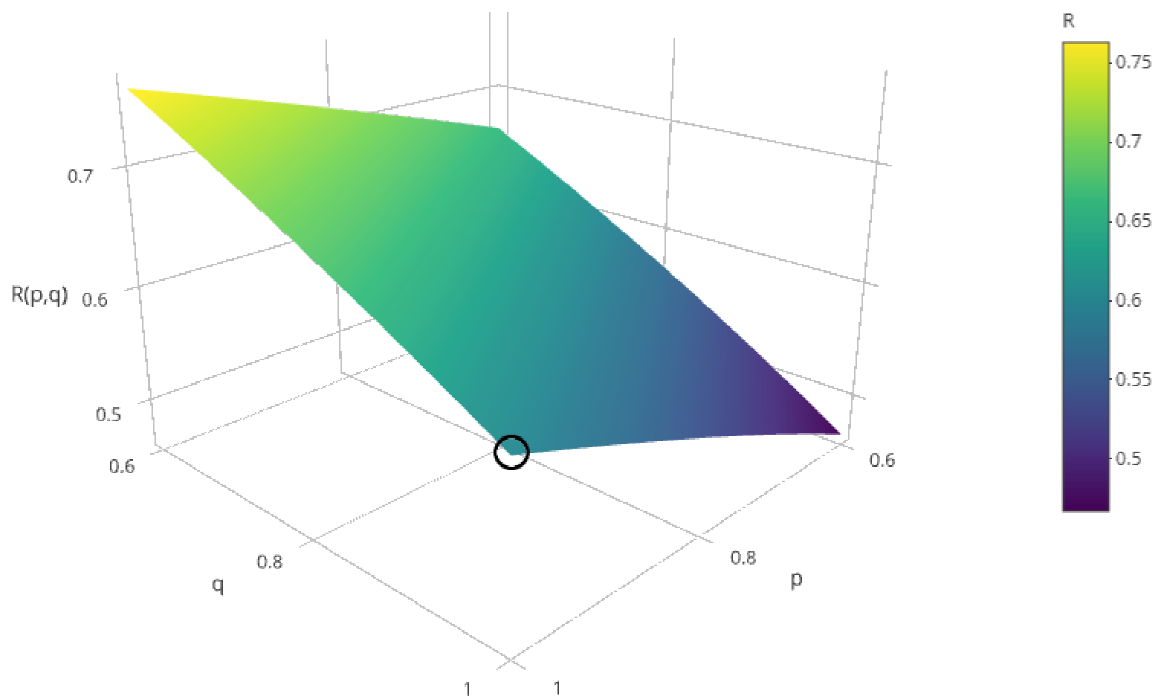
$$plim\left(T_G\right) = R\left(1,1\right) = \sqrt{\frac{plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right)}{plim\left(\frac{1}{n}\sum_i \widetilde{V}_i^2\right) + \left(\frac{\sigma_Y}{\psi}\right)^2}};$$

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

Figure 2: Example of Values of $plim\ T_G$



away from 1. Second, this point is even more salient when viewed relative to the magnitudes of the t-statistics we report in Section 6.2.3. 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 VR 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 VR services to DARS in 2000. We observe information on the VR services received along with some demographic and disability/health information. We decompose the sample into three large disability groups: people with

Table 1: Proportion of Clients Receiving Purchased Services of Each Type by Disability Group

	Cognitive Impairment	Mental Illness	Physical Impairment
Diagnosis & Evaluation	0.356	0.389	0.501
Training	0.407	0.292	0.131
Education	0.018	0.107	0.074
Restoration	0.208	0.282	0.380
Maintenance	0.301	0.298	0.182
Transportation	0.261	0.215	0.126
Other	0.029	0.057	0.060
# Observations	1009	1555	2612

Table 2: Selected Explanatory Variable Means Across Disability Groups

	Cognitive Impairment	Mental Illness	Physical Impairment
Male	0.506	0.404	0.475
White	0.557	0.710	0.647
Education	6.770	10.718	9.955
Special Education	0.332	0.025	0.014
Age	25.100	35.700	41.000
Transportation Available	0.460	0.741	0.824
Has Driving License	0.174	0.678	0.786
Musculo/Skeletal Disability	0.067	0.170	0.709
Internal Disability	0.061		0.337
Learning Disability	0.046	0.046	
Mental Illness	0.183		0.105
Significant Disability	0.573	0.619	0.606
Most Significant Disability	0.401	0.275	0.161

Note: Empty cells signify that the variable was not used for the particular disability group.

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.[17] In addition, we report the proportion of individuals who receive transportation assistance as part of the maintenance services category. There is significant variation across disability groups with respect to the reported proportions. For example, transportation services are received by 26.1% of clients with cognitive impairments, 21.5% of people with mental illness, and 12.6% of people with physical impairments.

Table 2 provides information about the means of selected explanatory variables for each of the three disability groups.[18] There are significant differences in means for most of the variables and in ways that make sense. For example, 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.[19] Having a driver's license is likely the reason individuals with mental illnesses or physical

[17]These types are described in more detail in Appendix 1 and in DPSS.

[18]Complete lists of explanatory variables, along with means and standard deviations, are available in DPSS.

[19]See, for example, Cyra, Mulroy, and Jans (1988), Stern (1993), Franklin and Niemeier (1998), and Bearse et al. (2004) for

Table 3: Labor Market Means Across Disability Groups

	Cognitive Impairment		Mental Illness		Physical Impairment	
	# Obs	Mean	# Obs	Mean	# Obs	Mean
Proportion Employed	58,522	0.253	90,190	0.301	140,418	0.335
log Quarterly Earnings	14,799	7.009	27,148	7.342	46,960	7.692

Note: log Quarterly Earnings are conditional on being employed.

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.[20] Table 3 provides means for the two dependent labor market variables, e_{it} (from equation (4)) and w_{it} (from equation (5)). 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[21] 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 passenger miles traveled (PMT), per capita vehicles operated in maximum service ($VOMS$), per capita vehicle revenue hours (VRH),[22] and per capita unlinked passenger trips (UPT).[23] For rural transportation agencies, due to the lack of data sources available,[24] we collected

descriptions of the transportation choice set of disabled people and how often they choose different alternatives.

[20]Details about these data are available in DPSS.

[21]Virginia, unlike all other states, has independent cities that are not part of their adjacent counties.

[22]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.

[23]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.

[24]Data for rural transportation agencies are collected only if they receive funding from the federal government. The U.S.

Table 4: Public Transportation Summary Statistics

	# Obs	Mean	Std Dev Across Counties	Std. Dev. Within Counties Across Time	Std Dev
Public Transportation Exists	133	0.542	0.398	0.308	0.498
PMT per Capita	39	56.190	47.550	30.373	48.256
VOMS per 1K Capita	41	0.266	0.130	0.095	0.149
VRH per Capita	41	0.602	0.353	0.191	0.385
UPT per Capita	93	5.992	9.050	5.078	11.082

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

	# Obs	R-Squared
Public Transportation Exists	3724	10.1%
PMT per Capita	629	6.9%
VOMS per 1K Capita	659	1.9%
VRH per Capita	657	2.8%
UPT per Capita	933	3.3%

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.[25] 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). Mattson (2015) suggests that *PMT*, *VRH*, and *UPT* are useful measures. We experiment with each because we do not know a priori which might be the most relevant 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.[26] 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).[27]

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

Department of Transportation started collecting information from such rural agencies only in 2007.

[25]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.

[26]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.

[27]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 6: Employment Effects of Service Receipt by Disability Group

	Cognitive Impairment		Mental Illness		Physical Impairment	
	Short Run	Long Run	Short Run	Long Run	Short Run	Long Run
Diagnosis & Evaluation	0.422	0.310	-0.228	-0.462	0.547	0.127
Training	0.694	0.458	0.631	0.541	0.324	0.408
Education	0.122	0.538	-0.299	-0.113	0.258	0.564
Restoration	-0.331	-0.458	-0.017	-0.127	0.515	0.357
Maintenance	-0.124	-0.059	0.054	-0.074	-0.007	-0.015
Other Services	0.383	0.419	0.128	0.049	0.385	0.269

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.

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^*_{it} 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 $\widehat{\alpha}^e_{j4} - \widehat{\alpha}^e_{j1} = 0.458$ 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.[28] Analogously, for short-run effects, $\widehat{\alpha}^e_{j3} - \widehat{\alpha}^e_{j1} = 0.694$ for the j corresponding to training services, suggesting a 27.8% 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

[28] 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).

Table 7: log Quarterly Earnings Effects of Service Receipt by Disability Group

	Cognitive Impairment		Mental Illness		Physical Impairment	
	Short Run	Long Run	Short Run	Long Run	Short Run	Long Run
Diagnosis & Evaluation	0.300	0.307	-0.085	0.032	0.163	0.318
Training	0.209	0.285	-0.055	0.136	0.009	0.172
Education	0.093	0.555	-0.085	0.146	0.318	0.364
Restoration	-0.308	-0.241	0.092	0.206	0.351	0.442
Maintenance	-0.251	-0.069	0.106	0.217	-0.165	0.029
Other Services	0.308	0.224	0.084	0.146	0.165	0.148

the individual might have difficulty working because she has no transportation available without the DARS transportation support.[29] We explore this issue more in Section 6.2.

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 $\widehat{\alpha}^w_{j4} - \widehat{\alpha}^w_{j1} = 0.172$; for the j corresponding to 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.[30] 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.[31]

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 Pseudo-LM Test Results for Interactions

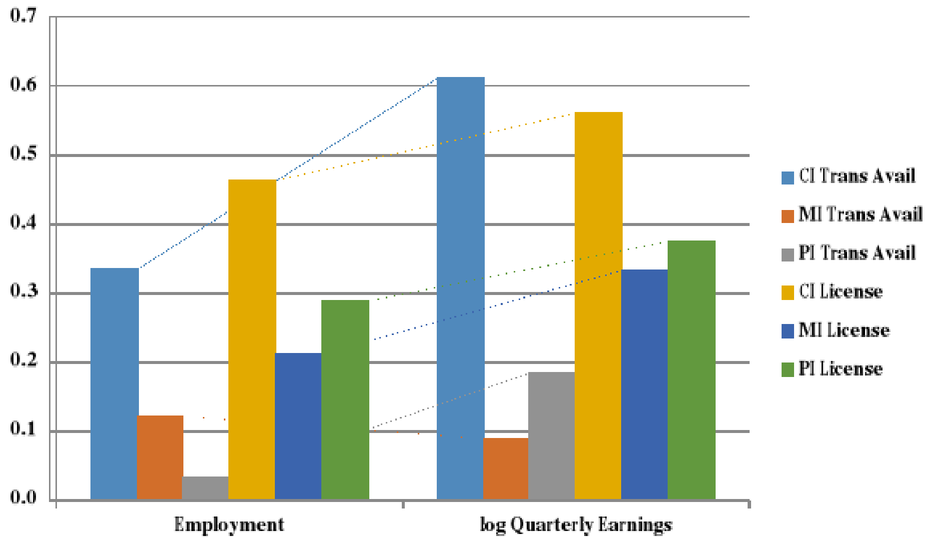
We now turn to the results of the our primary analysis, the pseudo-LM tests. In order to determine whether our estimates of the $\Delta\widehat{\rho}^k_{\tau n}$ statistics, defined in equation (8), indicate that the effect of public transportation is meaningful, we need a counterfactual against which to compare. For all of the tables of estimates that follow,

[29]This point was explained to us by Bob Schmidt.

[30]All estimates are statistically significant at the 5% level. No results for service choice values are reported here.

[31]Ermagun et al. (2016) use data that allow them to estimate the effect on public transportation availability on work. Their discussion of results excludes this effect which we interpret to mean that there was no statistically significant effect.

Figure 3: Transportation Effects on Labor Market Outcomes



we calculate estimates for two different groups of clients who received DARS services: the "beneficiary group" and the "placebo group."[32] We define the beneficiary group as those clients who required transportation services from the VR agency (and thus might benefit from having good paratransit because they lack an alternative).[33] Similarly, we define people in the placebo group as individuals who did not receive such services (and thus are unlikely have a need for better public transportation).[34] 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. 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.[35]

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 Table 9. Table 8 contains estimates of the long-run changes in correlations (i.e., the correlation more than eight quarters after service minus the correlation prior to service) between measures of the existence and intensity of public

[32] 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 paratransit (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 make use of paratransit because they have alternative means of transportation.

[33] If a VR client 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.

[34] We also have access to a question about whether a VR client believes she has access to transportation that could be used to define the groups. However, individuals may or may not consider paratransit as an option when they report that they "have no available transportation." To the extent that beneficiary and placebo groups based on that question would be contaminated, our results would be biased. Results from preliminary analyses based on group definitions using this question were very implausible: a high proportion of the results were negative. We know of no *theoretical* reason why (increased) access to transportation would reduce the beneficial impacts of VR service receipt on labor market outcomes. We believe that this "stated" group definition does not result in reasonable beneficiary and placebo groups, and prefer our "revealed" group definition instead.

[35] 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.

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

	Beneficiary Group	Placebo Group	Difference	t-Stat
Public Transportation Exists	-0.01	-0.03	0.02	1.08
PMT per Capita	0.07	-0.07	0.15	5.35**
VOMS per Capita	-0.03	-0.03	0.00	0.08
VRH per Capita	-0.01	-0.06	0.05	1.85*
UPT per Capita	0.06	-0.02	0.08	2.96**

Notes:
1) The beneficiary group contains clients who received transportation services from the VR agency, and the placebo group contains those who did not.
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 theoretical predictions, double-pound items are statistically significant at the 5% level, and single-pound items are statistically significant at the 10% level.

transportation in the community (discussed in Section 5.3) and the generalized residuals associated with being employed ($\Delta\hat{\rho}^{e}_{4n}$) for clients with cognitive impairments.[36] Each row contains separate estimates of the long-run change in the correlation of the residuals with the given public transportation measure. In the first column, the estimate for the beneficiary group of $\Delta\hat{\rho}^{e}_{4n} = -0.01$ (in the first row) is an estimate of the double difference of the correlation of unobserved components of employment and the existence of public transportation: a) the first difference is for people who received each VR service minus those who did not; and b) the second difference is for the correlation in the long-run quarters after service minus the correlation in the quarters prior to service. Taken alone, the point estimate of -0.01 would suggest that the interaction of the existence of public transportation and the receipt of VR services has a negative influence on long-run labor market outcomes. But, viewed relative to the estimate of -0.03 for people in the placebo group in the second column, the difference suggests that the interaction between VR services and existence of public transportation increases long-run employment probabilities for cognitively impaired clients.[37] The third column contains the estimate of this effect of 0.02, and the fourth column contains the t-statistic associated with the difference which is statistically insignificant for the existence measure. The estimate used here is akin to a differences-in-differences-in-differences estimator (e.g., Gruber, 2007). If significant, it would indicate that, relative to a baseline of individuals who did not receive transportation services from the VR agency, the existence of public transportation would, on average, improve the effectiveness of VR services on long-run labor market outcomes for cognitively impaired individuals who do not have other means of transportation available.

6.2.2 Discussion of Results

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.[38] Panel A (Panel B) reports the employment (conditional log quarterly earnings) estimates of the effect of the interaction of public transportation and VR service receipt.[39]

[36]Note that, in each specification reported, the existence 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.

[37]Since our estimates are based on fully identified structural models in DPSS, we are comfortable interpreting them as causal estimates. We explicitly note our choice of language so that the fact that they are correlations does not obscure their interpretation.

[38]As in Table 8, the beneficiary group is defined as clients who received transportation services from the VR agency.

[39]Note that the long-run employment results in the second column of Panel A of Table 9 are the same as the t-statistics in Table 8.

Table 9: Triple-Difference t-Statistics by Disability Group

Panel A: Employment Effects

	Cognitive Impairments		Mental Illness		Physical Impairment	
	Short-Run	Long-Run	Short-Run	Long-Run	Short-Run	Long-Run
Public Transportation Exists	0.96	1.08	0.95	-1.34	1.36	0.22
PMT per Capita	0.78	5.35**	0.75	1.33	2.24**	1.28
VOMS per Capita	0.54	0.08	0.89	-1.55	2.35**	0.59
VRH per Capita	0.81	1.85*	1.67*	1.10	3.36**	2.91**
UPT per Capita	0.10	2.96**	0.03	1.80*	2.37**	4.63**

Panel B: Conditional Log Quarterly Earnings Effects

	Cognitive Impairments		Mental Illness		Physical Impairment	
	Short-Run	Long-Run	Short-Run	Long-Run	Short-Run	Long-Run
Public Transportation Exists	-3.36##	-4.01##	0.18	0.84	-1.95#	-1.26
PMT per Capita	-0.76	2.93**	-1.34	0.56	-4.77##	-6.32##
VOMS per Capita	-0.24	2.40**	-0.38	0.96	-2.18##	-3.23##
VRH per Capita	0.12	1.70*	-0.32	1.65*	-2.47##	-3.96##
UPT per Capita	-0.46	1.66*	-0.71	1.76*	-3.55##	-3.77##

Notes:
1) 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.
2) For t-statistics that are inconsistent 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)), none of the short-run employment effects (Panel A) and only one of the short-run conditional earnings effects (Panel B) are statistically significant. Consistent with our expectations, however, we find evidence of positive, statistically significant effects in the long run for both employment and conditional earnings. Focusing on the intensive-margin effects (*PMT*, *VOMS*, *VRH*, and *UPT*), three of the four employment effects (*PMT*, *VRH*, and *UPT*) and all four earnings effects are positive and significant. Turning to the extensive-margin effect, the existence of public transportation employment effects are insignificant, but the analogous earnings effects are negative and significant (in both the short and long runs).

The overall pattern of results for cognitively impaired clients indicates that high-quality paratransit has a positive impact on both the employment and earnings outcomes of DARS clients in the long run. The most straightforward interpretation of these results is that effective paratransit provides access to employment opportunities that would not otherwise be available to VR clients with a more commute-constrained job search. Those additional opportunities both cause more clients with cognitive impairments to find a suitable match with an employer (increasing employment) and their matches to be of higher quality (increasing wages). Interpretation of the finding that the existence of paratransit reduces earnings conditional on being employed is less obvious. Given that the availability of paratransit does not affect employment and well-functioning paratransit increases earnings, the negative existence effects may be driven by poor-functioning systems that lower the earnings of VR clients, conditional on having a job. This is consistent with long or laborious commutes affecting productivity (reducing wages) or leading clients to work shorter days or fewer shifts (reducing hours).[40]

The estimates in columns (3) and (4) show that clients with a mental illness display few significant interactions between DARS service receipt and the public transportation variables. Only four of the 20 effects are significant. While all four significant effects are positive, none are significant at better than the 10% level. Taken as a whole, the results suggest that paratransit has, at best, a small effect on the impact that VR services have on the labor market outcomes of clients with mental illnesses.

The physical impairment interactions in columns (5) and (6) are the most consistent of any group. All four of the intensive-margin employment effects in Panel A indicate that the interaction of public transportation and DARS service receipt increases the probability of employment in the short run, and the *VRH* and *UPT* effects persist to the long run. In contrast, the conditional earnings measures in Panel B are predominantly negative.

Together the physical impairment results indicate that while public transportation may increase the effect VR service has on the probability of being employed, the interaction also decreases the earnings that DARS clients 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.[41]

Comparing the patterns of results across all three disability groups, we note that 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, but it is not clear why the results for clients with mental illnesses are so different from those for clients with cognitive impairments.

6.2.3 Discussion of Measures

We use the best available sources of paratransit data and augment them by hand where necessary, but we remind the reader that there are well-documented issues with the existing measures (discussed in section 5.3). One benefit of our research design is that we are able to experiment with multiple measures because we do not know a priori which is/are the most relevant ones in light of these issues. We next examine

[40] Phillips (2016) provides evidence from a correspondence study that employers favor low-wage employees with shorter commutes.

[41] 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.

whether and how these data issues are likely to affect our results. Then we assess the relevance of each of our measures.

There are two sources of mis-measurement in the data: a lack of information about quality and geographic imprecision. With regards to the former, we do not know the actual costs of using paratransit for each client who needs it; we just have county-wide measures of the total services used and provided that proxy for those costs. For example, we do not observe how long it takes an individual to travel via public transportation to the specific location of a potential job.[42] Alternatively, we do not know whether systems have onerous reservation or proof-of-eligibility requirements that render them impractical for some individuals. Our service provision and consumption measures are likely reasonable proxies for true costs, but we cannot rule out concerns. More popular systems (with large values of PMT and UPT) may have to coordinate more pick-ups and drop-offs that increase travel times, decrease reliability, and make the system more costly for individual users. Large $VOMS$ and VRH measures may be due to poorly scheduled or routed systems rather than robust ones.[43] This would mean that increases in our intensity measures reflect a more costly means of transportation, rather than a more efficient one, and would bias our estimates downward.

With regard to the second data issue (geographic imprecision), we do not actually know if public transportation is an available and feasible means of travel for each client in our data. We just know whether it exists in the individual's county. The public transportation system may serve a limited geographic area within the county (or may operate a restrictive number of hours each day). This mismeasurement would negatively bias both our existence estimates and the intensity measure estimates (that are conditional on availability).[44]

With these concerns in mind, we look at the results in Table 9 across all five of the paratransit measures used in our analysis to see if one stands out from the rest. Unfortunately, we are unable to pinpoint which individual measures of paratransit performance are most relevant. While none of the variables alone stand out, the intensive margin measures are frequently important predictors of the impact that paratransit has on VR client employment outcomes, but the existence measure results are less robust. This is consistent with an intuitive result: the existence of paratransit alone does not affect VR client employment outcomes, only well-functioning paratransit does. It is also consistent with the bias due to geographic imprecision being more severe than the bias caused by incomplete measures of costs. Section 4 included a discussion of testing in the presence of classical and nonclassical measurement error. Given the pattern of results in Table 9, we believe the more important source of measurement error in our data 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.[45] Thus, while we cannot rule out that measurement error explains the pattern of results we find, we prefer the more intuitive explanation that only a well-functioning paratransit system improves VR client outcomes.

6.3 Economic Significance

Since the pseudo-LM results do not have a real-world interpretation, we proceed by comparing the interaction effects in Table 9 to estimates from DPSS to provide a sense of economic significance relative to other interventions. Table 10 contain transportation variable effect estimates from both DPSS and our study (denoted as either DPSS or CSY in the first column of the tables). Panel A reports the effect of a change

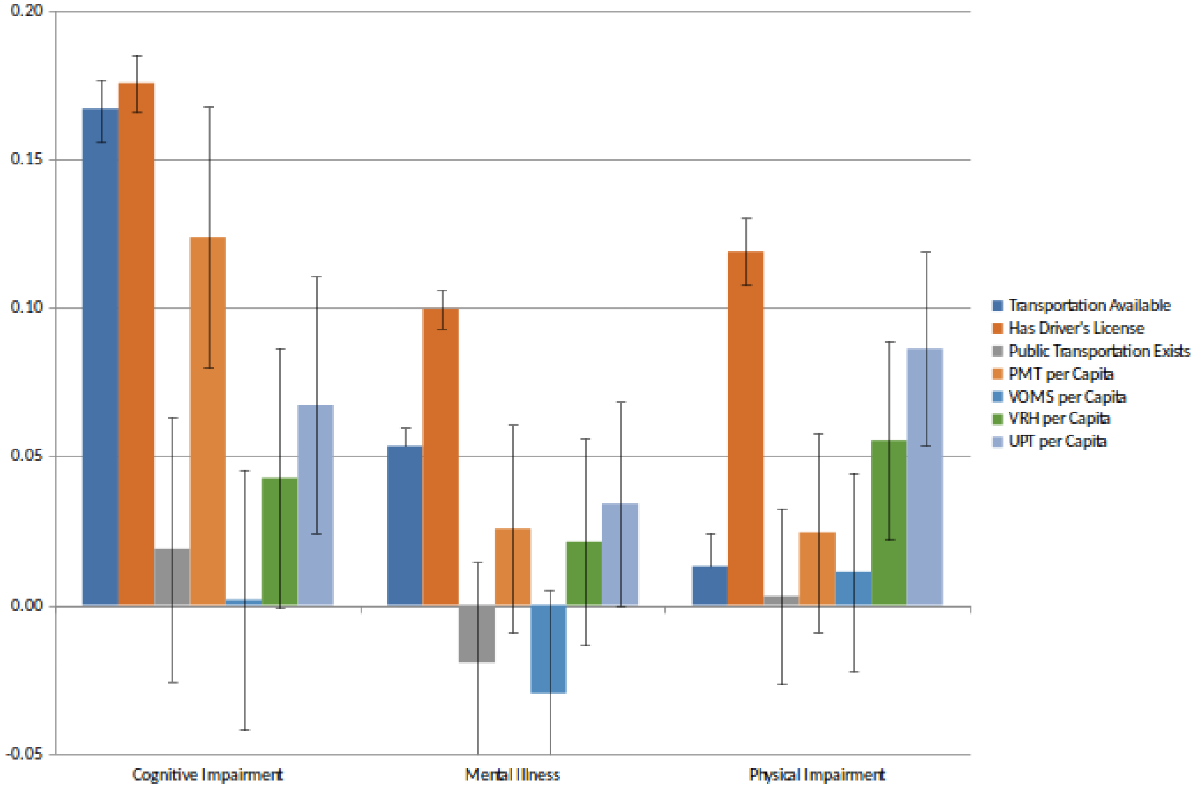
[42] While it is feasible to obtain reasonable measures of this time cost by modeling the transportation network (see Clapp, 2017), the lack of information on the home (and job) locations of individuals in our data (due to confidentiality reasons) prevents the use of this type of more detailed measure.

[43] 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.

[44] Our estimates are akin to an Intention to Treat (ITT) estimator, as all clients who are in need of transportation and have paratransit available in their community can be thought of as "treated," regardless of whether they actually make use of the paratransit system or not.

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

Figure 4: Employment Effects by Disability Group



in the given transportation measure on employment propensity, disaggregated by disability group. Panel B provides analogous estimates of the effects on conditional log quarterly earnings. To facilitate meaningful comparisons across transit measures based on different units, we report the effect of a one standard deviation increase in each given variable. The DPSS models do not allow transportation measures to have differential effects in the short and long run, so the table only contains one effect for *transportation available* and *has driver's license*. As described in Section 4, our estimates are calculated as the difference between the change in correlations for the beneficiary and placebo groups. The CSY effects are obtained by adjusting those double difference correlation estimates to allow for comparisons with the coefficient estimates in DPSS. The DPSS results are the same marginal effects as those previously reported in Figure 3. The CSY double difference estimates are obtained by multiplying correlations similar to those presented in Table 8 by a scaling factor that is easily calculated from sample statistics. The intuition behind the derivation of this scaling factor is that marginal effects are identified from the covariation of the given labor market outcome with the transportation measures (net of covariation with other explanatory variables), and our correlations are a function of that same covariance. Thus, we are able to derive a way to convert the latter measure into an approximation of the former. Note that estimates reported are approximations of the true marginal effects because they are based on derivatives of the objective function with respect to the paratransit measures which hold only locally. See Appendix 2 for full details of the adjustment process.

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 Panels A and B of Table 10 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 Panel A of Table 10 and the first grouping in Figure 4 show that three out of the five paratransit estimates (*PMT, VRH,* and *UPT*) indicate significant, positive employment effects that are of similar magnitudes to the transportation measures from

Table 10: Effects of Transportation Variables by Disability Group

Panel A: Employment Effects

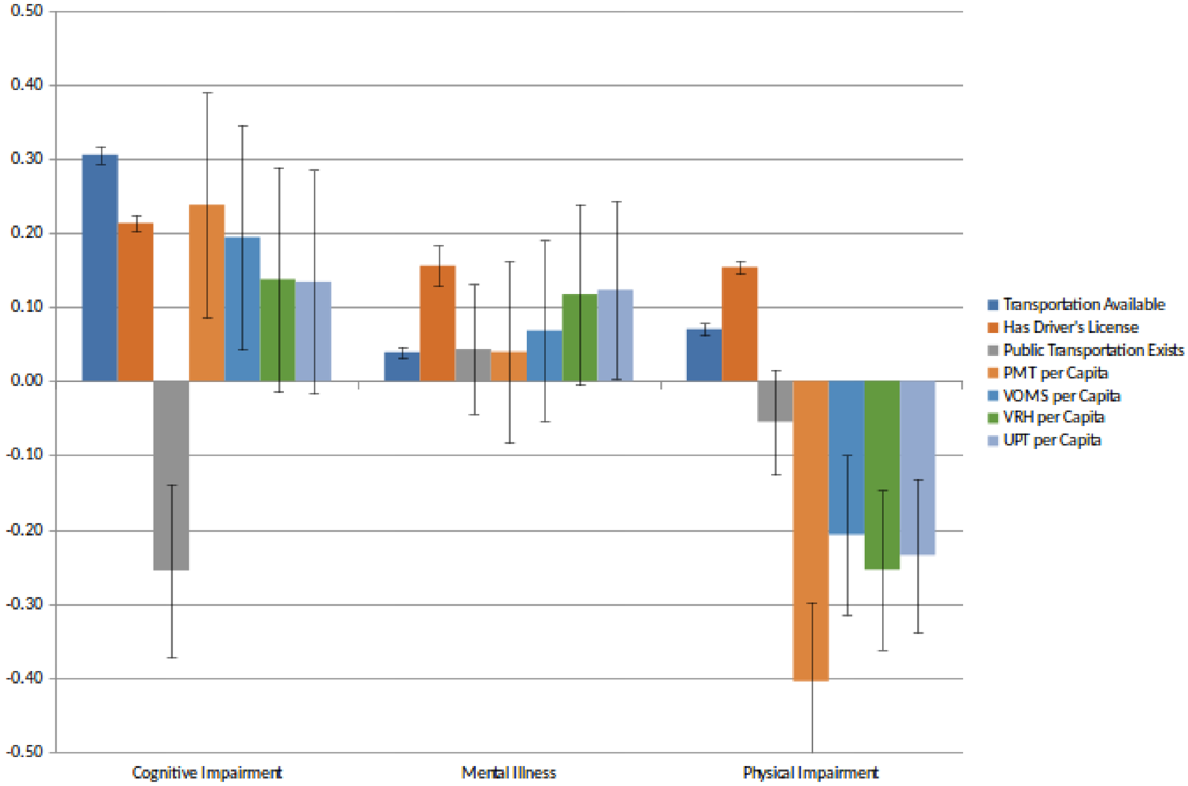
		Cognitive Impairments		Mental Illness		Physical Impairment	
		Short-Run	Long-Run	Short-Run	Long-Run	Short-Run	Long-Run
DPSS	Transportation Available		0.167**		0.053**		0.013**
	Has Driver's License		0.176**		0.100**		0.119**
CSY	Public Transportation Exists	0.024	0.019	0.020	-0.019	0.026	0.003
	PMT per Capita	0.026	0.124**	0.021	0.026	0.061**	0.024
	VOMS per Capita	0.018	0.002	0.025	-0.030	0.064**	0.011
	VRH per Capita	0.027	0.043*	0.046*	0.021	0.091**	0.055**
	UPT per Capita	0.003	0.067**	0.001	0.034*	0.064**	0.086**

Panel B: Conditional Log Quarterly Earnings Effects

		Cognitive Impairments		Mental Illness		Physical Impairment	
		Short-Run	Long-Run	Short-Run	Long-Run	Short-Run	Long-Run
DPSS	Transportation Available		0.305**		0.039**		0.070**
	Has Driver's License		0.213**		0.156**		0.154**
CSY	Public Transportation Exists	-0.266##	-0.256##	0.012	0.043	-0.119#	-0.055
	PMT per Capita	-0.077	0.238**	-0.123	0.039	-0.416##	-0.405##
	VOMS per Capita	-0.024	0.194**	-0.035	0.068	-0.191##	-0.207##
	VRH per Capita	0.012	0.137*	-0.029	0.117*	-0.216##	-0.254##
	UPT per Capita	-0.046	0.134*	-0.065	0.123*	-0.310##	-0.235##

Notes:
1) 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.
2) For t-statistics that are inconsistent theoretical predictions, double-pound items are statistically significant at the 5% level, and single-pound items are statistically significant at the 10% level.
3) The DPSS models do not allow transportation measures to have differential effects in the short and long run, so the table only contains one effect for transportation available and has driver's license.

Figure 5: Conditional log Quarterly Earnings Effects by Disability Group



DPSS (reporting that one had access to transportation or a driver's license).[46] Additionally, the "Long Run" column in Panel B of Table 10 and the first grouping in Figure 5 indicate that the long-run, earnings effects of all four of the intensive-margin measures (*PMT*, *VOMS*, *VRH*, and *UPT*) are both positive, significant, and of the same order of magnitude as the DPSS measures.[47] The existence measure is negative, but also of the same order of magnitude as the DPSS measures. The patterns suggest that, for people with cognitive impairment, effective 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 relatively larger than their paratransit analogs (which are generally not statistically significant). The conditional log quarterly earnings effects for the paratransit measures in Figure 5 are of much more similar magnitudes to the DPSS effects, but again, they are at best marginally significant. These results suggest that, for people with mental illness, improved paratransit is unlikely to have the same effect on long-run earnings outcomes as improved access to personal transportation.

Finally, the third grouping in each figure displays the relative magnitudes of the employment and conditional log earnings results for people with physical impairments. Panel A of Table 10 shows that people with physical impairments receive nontrivial, short-run employment benefits from improved paratransit, and there is evidence that those effects persist to the long run. Figure 4 shows that two of the five earnings effects (*VHR* and *UPT*) are statistically significant and positive in the long run. Both of these employment effects have magnitudes that are bracketed by the corresponding DPSS effects. Figure 5 shows that those gains are offset by robust, negative conditional earnings effects in the long run that are larger in magnitude than the personal transportation effects. The negative, statistically and economically significant earnings

[46] Although the 95% confidence interval for the *VRH* effect in Figure 5 contains zero, Table 10 indicates that the 90% confidence interval does not.

[47] *PMT* and *VOMS* are significant at the 5% level. *VRH* and *UPT* are significant at the 10% level.

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 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.[48] None of the *public transportation exists* effects are positive and statistically significant, but at least two system efficacy measures have a positive, significant, long-run effects on employment and/or earnings for two 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 with cognitive or physical impairments.

That these paratransit effects are not trivial in magnitude relative to those associated with measures of personal transportation is a striking result in and of itself for multiple reasons. First, 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, the use of personal transportation may be associated with less-severe limitations among people in the VR population. Finally, as has been established in the literature, the available data on paratransit service is lacking (Godavarthy et al., 2015). Despite our best efforts, better measures of the existence, efficacy, and use of paratransit are needed to more accurately measure the role that paratransit plays in the labor market outcomes of disabled people. All three of these facts suggest that the personal transportation effects should dominate their paratransit counterparts, but this is not the case. 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.

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

[48] 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.

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 several also persist to the long run.

There is also 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, particularly for those with a mental illness. 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, long-run 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 suggest that paratransit has a small, if any, affect on employment outcomes. We do not know why paratransit has such a limited impact on this group of VR clients.

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,[49]

$$\begin{aligned} \xi_i^* &= W_i\delta + \zeta_i, \; i = 1, 2, .., n, \\ \zeta_i &\sim iidN(0,1), \\ \xi_i &= \Upsilon(\xi_i^*) \end{aligned}$$

where ξ_i is observed, ξ_i^* is unobserved, and $\Upsilon(\xi_i^*)$ is a (possibly) nonlinear function of ξ_i^*. Equations (3) and (4) provide a binary choice example where $\Upsilon(\xi_i^*) = 1(\xi_i^* > 0)$.[50] Equation (5) provides a linear model example where, implicitly, $\Upsilon(\xi_i^*) = \xi_i^*$ and the restriction on the $Var(\zeta_i) = 1$ can be relaxed. Define the generalized residual (Gourieroux et al., 1987) for the model as

$$\widehat{\zeta}_i = E(\xi_i^* - W_i\delta \mid \xi_i).$$

Next, consider some other random variable

$$P_i \sim iid\left(\mu_p, \sigma_p^2\right)$$

with

$$E\left(\widehat{\zeta}_i \mid P_i\right) = \rho_{\widehat{\zeta}|P}(P_i - \mu_p), \tag{11}$$

at least locally, i.e., when $P_i - \mu_p$ is small.[51] 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

$$\widehat{\rho}_{\widehat{\zeta}|P} = \frac{n^{-1}\sum_i \widehat{\zeta}_i (P_i - \mu_p)}{n^{-1}\sum_i (P_i - \mu_p)^2} \tag{12}$$

be a consistent estimator of $\rho_{\widehat{\zeta}|P}$. Then, by definition, the estimated correlation of the generalized residuals and P is

$$\widehat{Corr}\left(\widehat{\zeta}, P\right) = \frac{n^{-1}\sum_i \widehat{\zeta}_i (P_i - \mu_p)}{\sqrt{\left[n^{-1}\sum_i \widehat{\zeta}_i^2\right]\left[n^{-1}\sum_i (P_i - \mu_p)^2\right]}}.$$

[49] The normality assumption is not necessary here and is made only for concreteness.

[50] Equations (1) and (2) also provide such an example, but they are not relevant to the exercise described in this appendix.

[51] This occurs globally for a large class of models where

$$\begin{pmatrix} \widehat{\zeta}_i \\ P_i \end{pmatrix} = \begin{pmatrix} 0 \\ \mu_p \end{pmatrix} + \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}$$

where $\nu = (\nu_1, \nu_2)'$ has a zero mean and a finite covariance matrix.

Using equation (12), we can write

$$\begin{aligned}
\widehat{\rho}_{\widehat{\zeta}|P} &= \frac{n^{-1}\sum_i \widehat{\zeta}_i \left(P_i - \mu_p\right)}{\sqrt{\left[n^{-1}\sum_i \left(P_i - \mu_p\right)^2\right]\left[n^{-1}\sum_i \widehat{\zeta}_i^2\right]}} \frac{\sqrt{n^{-1}\sum_i \widehat{\zeta}_i^2}}{\sqrt{n^{-1}\sum_i \left(P_i - \mu_p\right)^2}} \\
&= \widehat{Corr}\left(\widehat{\zeta}, P\right)\frac{\widehat{\sigma}_{\widehat{\zeta}}}{\widehat{\sigma}_p}.
\end{aligned}$$

Then, using estimates of $Corr\left(\widehat{\zeta}, P\right)$, $\sigma_{\widehat{\zeta}}$, and σ_p, we can evaluate $E\left(\widehat{\zeta}_i \mid P_i\right)$ from equation (11) and get an estimate of the marginal effect of P_i on ξ_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., $\left(\widehat{\zeta}_i, P_i\right) \sim iidN\left(0, \Psi\right)$ with covariance matrix Ψ.

References

[1] Amemiya, Takeshi (1981). "Qualitative Response Models: A Survey." *Journal of Economic Literature.* 19(4): 1483-1536.

[2] Arkansas Research and Training Center on Vocational Rehabilitation (1992). *President's Committee on Employment of People with Disabilities. Employment Priorities for the 90's for People with Disabilities.* Washington, DC.

[3] Ashenfelter, Orley (1978). "Estimating the Effect of Training Programs on Earnings." *Review of Economics and Statistics.* 60: 47-57.

[4] Bajari, Patrick and Matthew Kahn (2005). "Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities." *Journal of Business & Economic Statistics.* 23(1): 20–33.

[5] Bajari, Patrick and Matthew Kahn (2008). "Estimating Hedonic Models of Consumer Demand with an Application to Urban Sprawl." *Hedonic Methods in Housing Markets.* (eds.) A. Baranzini, J. Ramirez, C. Schaerer, and P. Thalmann. Springer, New York.

[6] Bayer, Patrick and Robert McMillan (2012). "Tiebout Sorting and Neighborhood Stratification." *Journal of Public Economics.* 96(11): 1129–1143.

[7] Bearse, Peter, Shiferaw Gurmu, Carol Rapaport, and Steven Stern (2004). "Estimating Disabled People's Demand for Specialized Transportation." *Transportation Research.* 38(9): 809-831.

[8] Black, Dan, Mark Berger, and Frank Scott (2000). "Bounding Parameter Estimates with Nonclassical Measurement Error." *Journal of the American Statistical Association.* 95(451): 739-748.

[9] Börsch-Supan, Axel and Vassilis Hajivassiliou (1992). "Health, Children, and Elderly Living Arrangements: A Multiperiod-Multinomial Probit Model with Unobserved Heterogeneity and Autocorrelated Errors." *Topics in the Economics of Aging.* (ed.) David Wise. Chicago and London: U. of Chicago Press, 79-104.

[10] CES, Inc. and TranSystems (2009). *JARC/New Freedom Program Performance Evaluation FY 2007 / FY 2008.* http://docplayer.net/15327024-Jarc-new-freedom-program-performance-evaluation.html.

[11] Checkovich, Tennille and Steven Stern (2002). "Shared Caregiving Responsibilities of Adult Siblings with Elderly Parents." *Journal of Human Resources.* 37(3): 441-478.

[12] Clapp, Christopher (2017). "Should My Car Move or Should I? A Model of Residential and Commuting Choices." Unpublished manuscript.

[13] Community Transportation Association of America (2016). http://web1.ctaa.org/webmodules/webarticles/anmviewer.as

[14] Cyra, D. M. Mulroy, and R. Jans (1988). "An Inventory of Twelve Paratransit Service Delivery Experiences." *Transportation Research Record.* 1170: 69–80.

[15] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2015). "The Effects of Vocational Rehabilitation for People with Cognitive Impairments." *International Economic Review.* 56(2): 399-426.

[16] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2017a). "The Effects of Vocational Rehabilitation Services for People with Mental Illness." *Journal of Human Resources.* Forthcoming.

[17] Dean, David, John Pepper, Robert Schmidt, and Steven Stern (2017b). "The Effects of Vocational Rehabilitation for People with Physical Impairments." *Journal of Human Capital.* forthcoming.

[18] Denson, C. (2015). "Public Sector Transportation for People with Disabilities: A Satisfaction Survey." *Journal of Rehabilitation.* 66(3): 29-37.

[19] Ellis, Elizabeth and Brian McCollom (2009). *TCRP Report 136: Guidebook for Rural Demand-Response Transportation: Measuring, Assessing, and Improving Performance.* Transit Cooperative Research Program, Washington, DC. Transportation Research Board.

[20] Ermagun, Alireza, Shaghayegh Hajivosough, Amir Samimi, and Taha Rashidi (2016). "A Joint Model for Trip Purpose and Escorting Patterns of the Disabled." *Travel Behaviour and Society.* 3: 51–58.

[21] Franklin, Joel and Debbie Niemeier (1998). "Discrete Choice Elasticities for Elderly and Disabled Travelers Between Fixed-Route Transit and Paratransit." *Transportation Research Record.* 1623: 31-36.

[22] Friedberg, Leora and Steven Stern (2014). "Marriage, Divorce, and Asymmetric Information." *International Economic Review.* 55(4): 1155-1199.

[23] Geweke, John (1988). "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference." *Journal of Econometrics.* 38(1/2): 73-89.

[24] Godavarthy, Ranjit, Jeremy Mattson, Del Peterson, and Jill Hough (2015). "Developing a Method for Assessing National Demand-Response Transit Level of Service." *Journal of Public Transportation.* 18(4): 1-15.

[25] Goodwill, Jay and Ann Joslin (2013). "Forecasting Paratransit Services Demand - Review and Recommendations." National Center for Transit Research, University of South Florida.

[26] Gourieroux, C., A. Monfort, E. Renault, and A. Trognon (1987). "Generalized Residuals." *Journal of Econometrics.* 34(1/2): 5-32.

[27] Greene, William (2009). "Discrete Choice Modeling." *The Handbook of Econometrics: Vol. 2, Applied Econometrics.* (eds.) T. Mills and K. Patterson. Palgrave, London.

[28] Gruber, Jonathan (2007). *Public Finance and Public Policy.* Worth Publishers, New York.

[29] Jolly, Debbie, Mark Priestley, and Bryan Matthews (2006). *Secondary Analysis of Existing Data on Disabled People's Use and Experiences of Public Transport in Great Britain.* Disability Rights Commission.

[30] KFH Group, Inc., Urbitran Associates, Inc., McCollom Management Consulting, Inc., and Cambridge Systematics, Inc. (2008). *TCRP Report 124: Guidebook for Measuring, Assessing, and Improving Performance of Demand-Response Transportation.* Transit Cooperative Research Program, Washington, DC: Transportation Research Board.

[31] Kittelson & Associates, Inc., Urbitran Associates, Inc., LKC Consulting Services, Inc., MORPACE International, Inc., Queensland University of Technology, and Yuko Nakanishi (2003). *TCRP Report 88:*

A Guidebook for Developing a Transit Performance-Measurement System. Transit Cooperative Research Program, Washington, DC: Transportation Research Board.

[32] Koffman, D., D. Lewis, D. Chia, J. Burkhardt, and M. Bradley (2007). *Improving ADA Complementary Paratransit Demand Estimation.* Transit Cooperative Research Program Report 119.

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

[34] Kreider, Brent and John Pepper (2007). "Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors." *Journal of the American Statistical Association.* 102(478): 432-441.

[35] Langer, Ashley and Clifford Winston (2008). "Toward a Comprehensive Assessment of Road Pricing Accounting for Land Use." *Brookings-Wharton Papers on Urban Affairs 2008.* (eds.) Gary Burtless and Jennifer Pack. The Brookings Institution Press, Washington, DC.

[36] Magill-Evans, J., N. Galambos, J. Darrah, and C. Nickerson (2008). "Predictors of Employment for Young Adults with Developmental Motor Disabilities." *Work.* 31(4): 433-442.

[37] Mattson, Jeremy (2015). *Rural Transit Fact Book 2015.* Small Urban and Rural Transit Center, Upper Great Plains Transportation Institute, North Dakota State University.

[38] 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.

[39] Mayock, Tom (2016). "Wages, Housing Prices and Commutes." *Real Estate Economics.* 44(1): 258-300.

[40] National Council on Disability. (2005). *The Current State of Transportation for People with Disabilities in the United States.* Washington, D.C.

[41] National Organization on Disability (2000). *N.O.D./Harris Survey of Americans with Disabilities, 2000.* http://www.nod.org/content.cfm?id=798.

[42] Phillips, David (2017). "Do Low-Wage Employers Discriminate Against Applicants with Long Commutes? Evidence from a Correspondence Experiment." Unpublished manuscript.

[43] Rosenbloom, S. (2007). "Transportation Patterns and Problems of People with Disabilities." *The Future of Disability in America.* (eds.) M. Field and A. Jette. Washington, D.C.: National Academic Press.

[44] Sabella, Scott, Jill Bezyak, and Robert Gattis (2016). "Barriers to Public Transportation: The Impact on Employment of Individuals with Disabilities." Unpublished manuscript.

[45] Sapper, Deborah, Jay Goodwill, and Holly Carapella (2009). *Impacts of More Rigorous ADA Paratransit Eligibility Assessments on Riders with Disabilities.* Florida Department of Transportation Research Center.

[46] Scheer, J., T. Kroll, T. Neri, and P. Beatty (2003). "Access Barriers for Persons with Disabilities: The Consumer's Perspective." *Journal of Disability Policy Studies.* 13(4): 221-230.

[47] Schmidt, M. and D. Smith (2007). "Individuals with Disabilities Perceptions on Preparedness for the Workforce and Factors that Limit Employment." *Work.* 28: 13–21.

[48] Stern, Steven (1993). "A Disaggregate Discrete Choice Model of Transportation Demand by Elderly and Disabled People in Rural Virginia." *Transportation Research, A-POL.* 27(4): 315-327.

[49] Stern, Steven (1997). "Simulation-Based Estimation." *Journal of Economic Literature.* 35(4): 2006-2039.

[50] Transit Cooperative Research Program (2013). *Methods for Forecasting Demand and Quantifying Need for Rural Passenger Transportation: Final Workbook, TCRP Report 161.* Washington, DC

[51] United States, Bureau of Economic Analysis (2016). *CA1 Personal Income Summary: Personal Income, Population, Per Capita Personal Income.*

[52] United States, Bureau of Transportation Statistics (2002). *Omnibus Survey.* https://www.apps.bts.gov/programs/omnibus_surveys/targeted_survey/2002_national_transportation_availability_and_use_survey/public_use_data_files.

[53] United States, Bureau of Transportation Statistics (2003). *National Transportation Statistics 2003, Table 1-8, ADA Lift- or Ramp-Equipped Transit Buses.* http://www.bts.gov/publications/national_transportation_statistics/ 2003/html/table_01_08.html.

[54] United States, Department of Education (2001). *RSA Program Data and Statistics: FY 2001*, Table 16. http://www2.ed.gov/rschstat/eval/rehab/statistics-2001.html.

[55] United States, Department of Transportation (2016). *TS2.1 - Service Data and Operating Expenses Time-Series by Mode.* www.transit.dot.gov/ntd/data-product/ts21-service-data-and-operating-expenses-time-series-mode.

[56] United States, General Accounting Office. (2003). *Transportation – Disadvantaged Populations: Some Coordination Efforts among Programs Providing Transportation Services, but Obstacles Persist.* http://www.gao.gov/new.items/d03697.pdf.

[57] United States, General Accounting Office (2012). "ADA Paratransit Services: Demand Has Increased, but Little is Known about Compliance." GAO-13-17.

[58] Virginia Department of Rail and Public Transportation (2016). *Transit Development Plans.* http://drpt.virginia.gov/transit/major-transit-initiatives/major-transit-planning/transit-development-plans/.

[59] Weller, B. (1994). "Unmet Needs for Developmental Disabilities Services." *Population and Environment.* 15(4): 279-302.

[60] West, M., T. Hock, K. Wittig, and V. Dowdy (1998). "Getting to Work: Training and Support for Transportation Needs." *Journal of Vocational Rehabilitation.* 10(2): 159-167.