Investigating Spatial Mismatch in the District of Columbia-Maryland-Virginia Metropolitan Region: Pre- and Post-Analysis Plan

Scott Brown, Hannah Hill, Peyton Weber

8 May 2023

1 Introduction

"Spatial mismatch" refers to the discrepancy between the locations of suitable job opportunities and job seekers' residences and neighborhoods. Spatial mismatch can create significant challenges for residents belonging to metropolitan areas that consist of sprawling suburbs. In this project, we seek to deepen the field's understanding of spatial mismatch in the geospatial context of the DMV region. Specifically, (i) does increasing accessibility to the WMATA metro via new Purple Line stations increase likelihood of successful employment outcomes? (ii) what is the impact of shifting proximity (in miles) to new Purple Line stations on employment outcomes?

We will employ a difference-in-difference model in neighborhoods surrounding designated new Purple Line stations in various radial groups to investigate spatial mismatch. Various studies have found that greater mismatch exacerbates regional unemployment rates and persistent joblessness. As a consequence, affected communities may likely experience higher rates of crime and social isolation or exclusion among affected individuals.The construction and expansion of public transit is one popular approach to mitigate spatial mismatch and improve labor market outcomes. We will study the direct impacts of this intervention on employment outcomes through this natural experiment.

2 Motivation and Conceptual Framework

2.1 Motivation

This proposed analysis of the Purple Line is grounded in previous studies of public transit projects in other geographies. Fan et al. (2012) examined the impact of a new light-rail line in the Minneapolis metropolitan area on access to jobs for high-, medium-, and low-wage workers; they found positive effects on job accessibility for all cohorts. Holzer et al. (2003) leveraged the construction of a new heavy-rail system in the San Francisco Bay Area to conduct a survey of firms; they found increased hiring of residents near new stations.

While many studies use administrative data to draw conclusions on new metro stations' impacts, our proposed analysis would include both administrative and longitudinal survey data. The primary motivation for conducting surveys is that we can collect unique data points during pre- and post-treatment periods. After-the-fact analysis of administrative data limits possible lines of inquiry. The delayed construction and opening of the Purple Line in 2026 represents a

unique opportunity for a natural experiment. Holzer et al.'s (2003) analysis of the new metro line in the San Francisco Bay Area offers a useful model for pre- and post-treatment surveys of firms before and after the opening of a new heavy-rail line to examine their propensity to hire workers residing near new stations. A similar analysis of the DC area labor market offers an opportunity to address previous limitations, such as restrictions to arbitrary distance inclusion criteria and the exclusion of jobseekers.

2.2 Conceptual Framework

Spatial mismatch is an issue that labor market studies (and policy practitioners) have yet to resolve. This market failure occurs when there are unequal opportunities for employment based on where an individual resides. In other words, geographical and transportation barriers prevent transactions in the labor market, i.e., the market is not perfectly competitive because qualified applicants and hiring firms cannot transact. Transportation challenges prevent jobseekers from gaining employment in distant employment hubs, and hiring firms may choose to reject applicants from particular geographies.

We hypothesize that this market failure is pronounced for poor households and single- and zero-car households for whom a new metro line would enable more hiring opportunities. We choose to examine the market failure among Maryland's workers as they face above-average unemployment rates. Even when access to public transportation has no effect on the rate at which a company hires minority applications, spatial mismatch implies that companies that want to avoid hiring certain or all minority applicants may choose to establish their firms in publicly inaccessible locations.

3 Sample and Intervention

3.1 Purple Train Stations

Our natural experiment takes place in neighborhoods between the District of Columbia and Maryland along the Bethesda-New Carrollton corridor. These neighborhoods vary greatly in terms of socioeconomic status, racial and ethnic demographics, commuting tendencies, housing markets, and resource accessibility. Because we are investigating the impact of *new* stations, we will exclude the following WMATA stations that are already built and service the surrounding neighborhoods: Bethesda, Silver Spring, College Park-UMD, and New Carrollton. The remaining station sample size is 17, for the following stations:

- Connecticut Avenue
- Lyttonsville
- 16th Street-Woodside
- Silver Spring Library
- Dale Drive
- Manchester Place
- Long Branch
- Piney Branch Road
- Takoma/Langley
- Riggs Road
- Adelphi Road-UMGC-UMD
- Campus Drive-UMD
- Baltimore Avenue-UMD
- Riverdale Park North-UMD
- Riverdale Park-Kenilworth
- Beacon Heights-East Pines
- Glenridge

Through these 17 stations, we will incorporate a within-station design to the difference-in-difference model in order to control for unobserved heterogeneity that may impact employment and job-seeking outcomes. For example, the new stations that connect various parts of the University of Maryland campus together may consist primarily of working aged-adults but are instead full-time students. These individuals who use the UMD stations as their home station thus may have entirely different characteristics than individuals of the same age but use other locations along the Purple Line as their home station. The within-station design accounts for this unobserved heterogeneity while observing the changes in outcomes before and after the treatment (reduction in travel time) is implemented.

3.1.2 Individuals

The ideal sample size of individuals for each station is 250, summing 4,250 total individuals throughout the experiment. Our target population consists of working-age adults aged 18-65 years old. We will utilize public record lists of residential addresses to randomly extend invitations to participate with a voucher incentive through MDOT and/or WMATA. Additionally, we will disseminate our survey at the 17 Purple Line stations listed above to widen the pool of potential participants and increase the likelihood to meet our desired goal of 250 individuals per station.

We define our radial groups as follows: the first difference-in-difference trial will consist of a treatment group that lives one quarter of a mile away from a planned, purple line metro stop, and the control group will comprise those who live one half of a mile away from a planned, purple line metro stop. The second difference-in-difference trial treatment group will consist of workers one half of a mile away compared to workers one mile away from a planned, purple line metro stop. Finally, the third trial will compare those defined by living one full mile away from a planned metro stop to those living five miles away from a planned metro stop.

3.2 Intervention

Our natural experiment allows us to evaluate a direct, small-scale intervention: a public-private partnership with the Maryland Department of Transportation (MDOT) and Purple Line Transit

Partners, LLC. Through the partnership, the state of Maryland seeks to correct severe traffic congestion in the Beltway (the highway that encircles the District of Columbia) while improving mobility for households who are dependent on public transportation. Current projections estimate that roughly 17,000 daily auto trips will be eliminated while reducing travel times by nearly 40%. Our model tests whether expanding access to public transportation for low-income households who tend to live farther away from job opportunities not only improves efficiency of the overall public transportation system, but also improves socio-economic equity and advancement as well as reduce negative externalities. Testable implications of results include proximity to transit stops on unemployment rates and employment patterns. In the DMV, it is possible that the additional transit stops have no effect on the mobility and thus access to employment opportunities of any group of individuals.

3.2.1 Surveys

We intend for our survey to include questions related to age, race, income, home sales, car ownership, employment, political ideology, and residence information. We would like to collect demographic data on our participants to control for important, potentially confounding factors that could relate to *both* the main independent variable (reduction in travel time), as well as the outcome variable (employment opportunities). Crucially, we would additionally survey respondents on their sentiments regarding the extent to which their employment opportunities either have or have not expanded. Using the interviewees' responses from the pre and post periods, we will then create an indicator variable to represent the outcome variable of interest - employment opportunity. The dummy variable will be coded equal to 0 if the observed participant did *not* express improved employment opportunities and coded equal to 1 if the participant expressed sentiments of improved employment opportunities from the first survey to the second survey administered.

3.2.2 Vouchers as incentives

Vouchers, subsidized by the MDOT and WMATA, will be distributed to all participants who successfully complete the survey. Participants will have two opportunities to receive vouchers; once when they complete the pre-survey before Purple Line stations have officially opened and once when they complete post-survey one year later. Vouchers will be disseminated either through a physical WMATA card usable at all WMATA and Purple Line stations or through a participant's existing digital WMATA card on their phone, as indicated by the participant's preference on the surveys.

3.3 Randomization

The treatment in our quasi-experimental design is defined by reduction in travel time caused by relative proximity to the new metro stations. The outcome variable of interest in an indicator variable that represents employment opportunity. Those that comprise the treatment groups will be individuals who meet the radial group characteristics defined above. Similarly, the control groups will comprise workers who are not receiving the treatment of relatively closer metro

accessibility. We rely upon the difference-in-difference model, because we want to estimate the effect of the specific natural intervention of transit expansion. We will compare the changes in employment opportunities for the treated workers with the changes in employment opportunities for the non-treated workers over the course of the metro project construction.

3.4 Risks

Because this is a natural experiment, there are no obvious undue risks that this study may impose on participants. One potential concern is that participants may distribute their WMATA voucher funds to other individuals outside of the natural experiment, however, these funds are just incentives to participate and are not actual intervention mechanisms.

3.5 Data sources

This study employs a two-wave longitudinal survey over the course of one year. The first-wave of data collection will be conducted in the two weeks prior to completion of the Purple Line extension project. The second-wave of data collection will be conducted approximately one year following the project completion. Both surveys will collect the same data through the same exact questions outlined in Section 3.2.1 using the same collection methods of either a digital survey completed online or on a smartphone or through mailing in a paper survey or handing in the paper survey to designated MDOT/WMATA workers at any participating stations.

We would access recent Public-Use Microdata Sample (PUMS) data to examine the following variables:

- Commute times
- Employment
- Residence
- Car-ownership
- Home sales

3.6 Timeline

- Pre-survey of individuals (0-2 weeks before Purple Line stations officially open)
- Distribution of pre-survey voucher incentives (immediately upon survey completion)
- Post-survey of individuals (52-54 weeks after Purple Line stations officially open)
- Distribution of post-survey voucher incentives (immediately upon survey completion)

4 Hypotheses and Outcomes

4.1 Hypotheses

The null hypothesis is that the change in the employment opportunities for the treatment group will be exactly the same as the change in the employment opportunities for the control group. The alternative hypothesis would be that the change in job opportunities for the treatment group, both before and after the public transit expansion, will **not** be the same as the change in the job opportunities for the control group over the same duration. We are interested in testing the null

hypothesis that workers who have access to the metro **do not** receive any additional gains in hiring opportunities compared to those who do not have metro accessibility. If our study's conclusions find that we may reject this null hypothesis, we are additionally interested in measuring the size of these potential gains. We must assume stable, comparable trends between treatment and control groups that would have persisted if the Purple line project were not to be completed. The overall change in employment for metro inaccessible workers subtracted from the overall change in employment for metro accessible workers will be our difference estimate.

4.2 Primary outcomes

We hypothesize that our difference estimate will allow us to reject the null hypothesis. Workers who reside close to the planned, purple line metro stations will receive access to quality jobs, expanding their employment probabilities, that were not available to the workers of interest prior to the transit expansion. Should our study's hypothesis prove accurate, this would have tremendous positive implications for future metrorail expansion projects.

5 Estimation methodology

For ease of interpretability, we intend to use a linear probability model to measure employment opportunities for our difference-in-difference technique. Because this technique is typically applied in a regression model in which indicator variables are used for time and the treatment effect, we propose the following model for our study:

$$Y= \beta0 + \beta1*[Time] + \beta2*[Intervention] + \beta3*[Time*Intervention] + \beta4*[Covariates]+\varepsilon$$

Relying upon our survey data, we will create a dummy variable named "treatment" that represents whether an observation belongs to the treatment or control groups. We will similarly create a time dummy variable that represents if the outcome is being measured in the pre-expansion period or in the post-expansion period. Finally, as stated previously, we will create an indicator variable for our outcome variable of interest. The dependent variable will serve as a representation of whether or not an observation expressed having high employment probability.

5.1 Limitations

Due to this intervention being a natural experiment, there are several limitations in regards to collecting the ideal sample size and randomizing participants into the study.

5.1.1 Attrition and survey non-response

Because we propose a longitudinal study, our ability to draw conclusions will depend on the rate of attrition between pre- and post-treatment measurements. In survey studies that last over years, we will undoubtedly lose respondents as they, for example, change addresses or otherwise decide to not respond. Attrition also introduces bias, as attrition is not random, e.g., low-income people

whose rent increase required them to move may be more likely to attrite as they change addresses.

At the same time, attrition and survey non-response may make it challenging to compile a large enough sample stratified by important characteristics like race. In other words, the people that decide to enroll in the study may systematically vary from the rest of the target population. If self-selected survey participants have different job-seeking behaviors, our estimated impact may only represent them, not the rest of the population. In short, self-selection bias in this particular study may impose limitations on our ability to draw meaningful and generalizable conclusions.

5.1.2 Treatment and control groups

Another limitation is introduced by our choice of radial groups and the risk of fuzzy treatment and control groups. Just because two households may live at the same distance to the new station as the crow flies, they could live in very different built environments and their actual commutes to the same station may look very different. Perhaps one commuter would experience worse traffic, or has to take a longer route around a park. Radial distance doesn't reflect these issues, meaning that our division of treatment and control will not be perfectly discrete.

One potential workaround to this is using estimated driving or walking times from every given point, but the labor required to do this would complicate measurement and potentially increase project costs. All that said, radial distance should serve as a useful proxy for access to the Purple Line stations. And we importantly chose to measure multiple potential radial groups so that we can find the most meaningful division between treatment and control based on real-world commuting and employment behaviors.

5.1.3 Conceptual limitations

Research design challenges aside, there are other conceptual limitations to this proposed study. The primary limitation of this design is that results will reflect only a particular snapshot in time; it's difficult to anticipate how long it would take for the Purple Line stations to have their full impact on the labor market. Workers won't likely quit their job immediately and opt for a commute newly enabled by the Purple Line. It may take several years for such changes to occur. This study, unless stretched out over multiple post-treatment surveys, won't capture this impact and therefore may underestimate the Purple Line's effects.

Moreover, surveys and administrative data will not capture the in- and out-migration in the region affected by the construction of the Purple. While the study will offer insight regarding longer-term residents, the Purple Line will undoubtedly attract new residents that will not be included in the study. In addition to spurring new in-migration, the Purple Line will have other spillover effects (for example in property value, amenities, etc.) that will not be assessed in the study.. In short, while our study aims to quantify certain economic impacts of public infrastructure projects, it will not offer an end-all, be-all measurement.

References

Andersson, Fredrik, John C. Haltiwanger, Mark J. Kutzbach, Henry O Pollakowski, and Daniel H. Weinberg. "Job Displacement and the Duration of Joblessness: The Role of Spatial Mismatch." *National Bureau of Economic Research* Working Paper Series, no. 20066 (April 2014): http://www.nber.org/papers/w20066.

Chung, Payton. "Jobs Are Clustering in Parts of the Region, but the East Is Falling Behind." Accessed May 8, 2023. https://ggwash.org/view/37817/jobs-are-clustering-in-parts-of-the-region-but-the-east-is-falling-behind.

Columbia University Mailman School of Public Health. "Difference-in-Difference Estimation." Last modified August 3, 2016. https://www.publichealth.columbia.edu/research/population-health-methods/difference-difference-estimation.

Fan, Yingling, Andrew E Guthrie, and David M Levinson. "Impact of Light Rail Implementation on Labor Market Accessibility: A Transportation Equity Perspective". *Journal of Transport and Land Use* 5, no. 3 (Dec. 2012): 28-39. https://doi.org/10.5198/jtlu.v5i3.240.

Hellerstein JK, Neumark D, McInerney M. "Spatial Mismatch Or Racial Mismatch?" *Journal of Urban Economics* 64, no. 2 (Sept. 2008): 464-479. Doi: 10.1016/j.jue.2008.04.003. PMID: 19727422; PMCID: PMC2614276.

Holmes, Elizabeth and Natalie Kneebone. "The Growing Distance between People and Jobs in Metropolitan America." *Brookings*. Last modified March 24, 2015. https://www.brookings.edu/research/the-growing-distance-between-people-and-jobs-in-metropolitan-america/

Holzer, Harry J., John M. Quigley, and Steven Raphael. "Public Transit and the Spatial Distribution of Minority Employment: Evidence from a Natural Experiment." *Journal of Policy Analysis and Management* 22, no. 3 (2003): 415–41. http://www.jstor.org/stable/3326034.