Title

Do Public Transit Investments Promote Urban Economic Development? Evidence from Bogotá, Colombia

Authors

Heres, David R.
Jack, Darby
Salon, Deborah

Publication Date

2013-02-01

Research Report – UCD-ITS-RR-13-01

Do Public Transit Investments Promote Urban Economic Development? Evidence from Bogotá, Colombia

February 2013

David R. Heres
Darby Jack
Deborah Salon

Institute of Transportation Studies ◦ University of California, Davis
One Shields Avenue ◦ Davis, California 95616
PHONE (530) 752-6548 ◦ FAX (530) 752-6572
www.its.ucdavis.edu

Do public transit investments promote urban economic development? Evidence from Bogotá, Colombia

David R. Heres[a], Darby Jack[b], and Deborah Salon*[c]

Abstract: In 2000, the city of Bogotá, Colombia embarked on a grand land use and transportation system experiment. The transformation of Bogotá included building the TransMilenio Bus Rapid Transit (BRT) system, a city-wide system that offers speed and convenience similar to that of an underground metro. TransMilenio is widely regarded as a success, and cities around the world are planning or building similar systems.

In this paper, we use a repeated cross-section labor market dataset to assess whether access to the new BRT system affects the incomes of those who live in station area neighborhoods. Our results indicate that the opening of the TransMilenio system was associated with increased income for those living near – but not immediately adjacent to – trunk line stations. This relationship is strongest in the lower and middle-income range. There are at least two possible explanations for this result: 1. existing residents earn higher wages, or 2. higher income workers move to the neighborhood. Our data do not allow us to distinguish clearly between them, but available evidence suggests that much of the effect is likely due to relocation. Our results stand in contrast to prior work, which has largely suggested that improvements in public transit will tend to reduce wages in station areas.

KEYWORDS: income; Bus Rapid Transit; spatial analysis

Notes: 1. The three authors of this paper share equal credit and equal responsibility for the work presented herein. The order of authorship listed is simply according to the order in which our surnames appear in the alphabet. 2. Earlier versions of this work were presented at two professional conferences: the 2009 Association of Collegiate Schools of Planning Annual Conference and the 2010 Kuhmo-Nectar Conference on Transport Economics.

[a] Centro de Investigacion y Docencia Economicas, Mexico City, Mexico
[b] Mailman School of Public Health, Columbia University, New York, USA
[c] Institute of Transportation Studies, University of California, Davis, CA, USA
email: ddsalon@ucdavis.edu
tel: 530-752-6548
*corresponding author

1. Introduction

In 2000, Bogotá embarked on a grand land use and transportation system experiment. Bogotá had about 7 million inhabitants, and the unemployment level was approximately 10 percent. Automobiles dominated the city, though only 19 percent of the population lived in car-owning households. Traffic congestion was severe, 70 percent of the particulate matter that clouded the air came from tailpipes, pedestrian safety was compromised, and parked cars clogged even the sidewalks (Echeverry et. al. 2005). The city government, under then-mayor Enrique Peñalosa, implemented a series of new policies regarding the use of public space and began to make substantial investments in public infrastructure - all of which aimed to increase the standard of car-free living for Bogotanos. These policies and investments have continued under subsequent city governments, and urban planners now cite Bogotá as a model city.

The specifics of the transformation of Bogotá include taking back the sidewalks for people, building approximately 350 kilometers of bicycle paths – many of them through poor neighborhoods – and building the TransMilenio Bus Rapid Transit (BRT) system. TransMilenio is a city-wide system that offers speed and convenience similar to that of an underground metro. Buses run in dedicated lanes and riders purchase tickets as they enter covered bus stations. The envisioned system is huge – with 400 kilometers of dedicated trunk routes plus feeder buses – but is not yet complete. The first lines opened in December of 2000, and additional lines opened each year through 2006. Figure 1 provides a map of the evolution of the system. Currently, there are over 80 kilometers of dedicated busways and approximately 500 kilometers of feeder bus routes. After only 6 years of operation, the system moved more than a million passengers each weekday (Cain et. al. 2006). Data from August 2012 put average weekday system ridership at 1.75 million daily (www.transmilenio.gov.co).

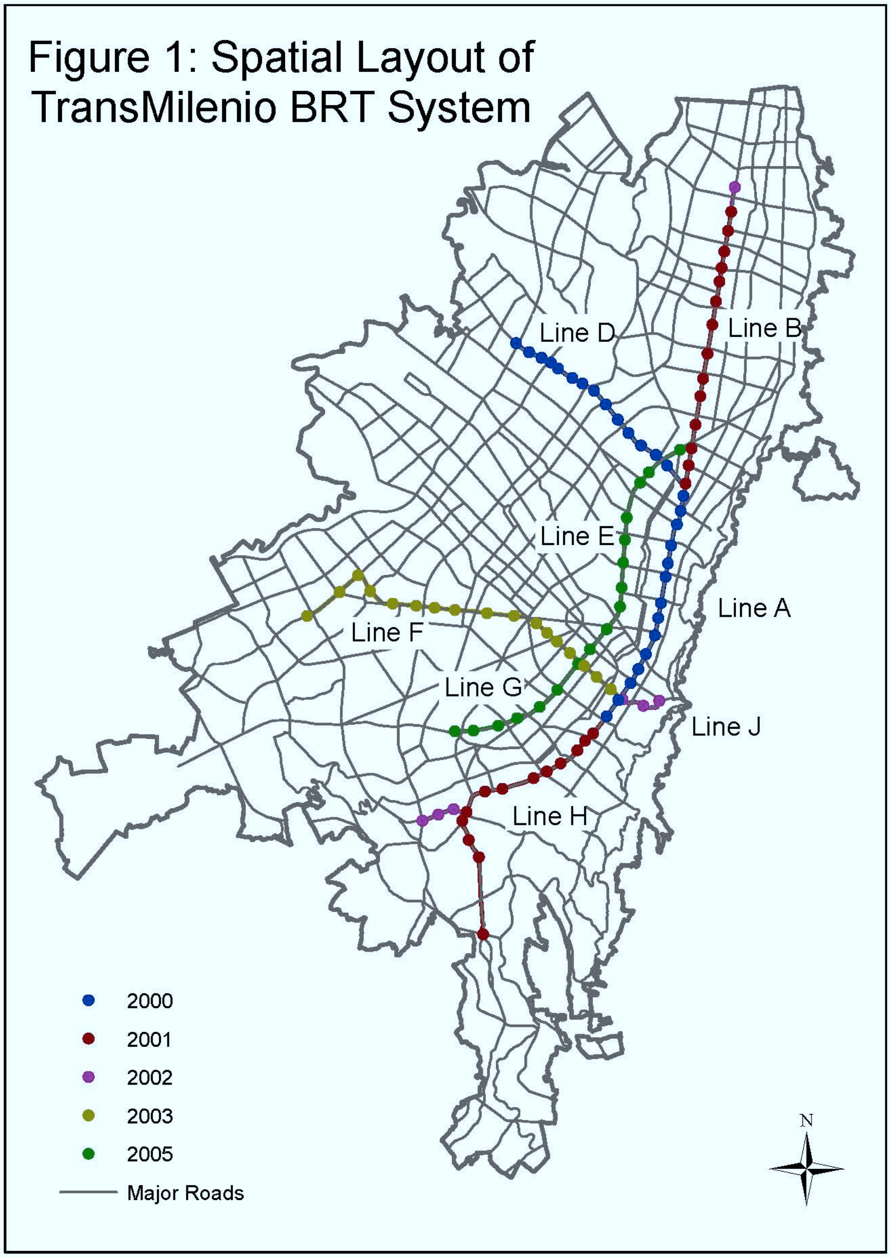

Figure 1: Spatial Layout of TransMilenio BRT System

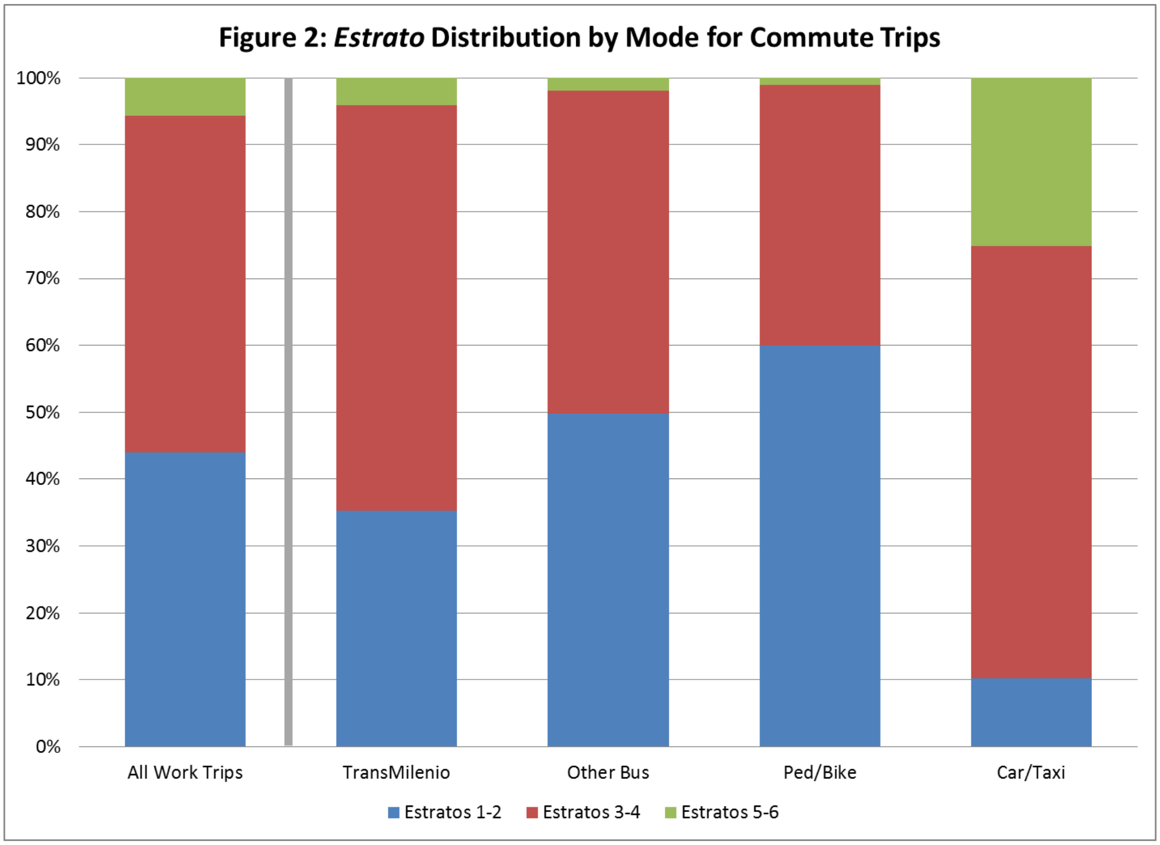
Figure 2: *Estrato* Distribution by Mode for Commute Trips
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
All Work Trips
TransMilenio
Other Bus
Ped/Bike
Car/Taxi
Estratos 1-2
Estratos 3-4
Estratos 5-6

Based on the 2005 Bogotá travel survey (*Encuesta de Movilidad*), TransMilenio carries over 10% of eligible[1] commute trips overall, and approximately 17% of commute trips for those living within walking distance of a station.[2] TransMilenio carries a similar percentage of non-work trips as well. TransMilenio commuters are slightly wealthier than riders of Bogota's conventional buses (see Figure 2).

TransMilenio has been cited by city planners and transportation engineers around the world as a success. Transit travel times have declined substantially, transit-related accidents have plummeted along TransMilenio routes, and there is even a measurable decrease in air pollution along TransMilenio corridors (Echeverry et. al. 2005). Inspired in part by the success of the TransMilenio, BRT systems have multiplied rapidly in cities around the world. As TransMilenio-style transit systems become increasingly common, understanding the full impact of these systems on urban economies becomes increasingly important.

Transit advocates often argue that good transit systems promote urban economic development by improving job matching between employers and workers who do not own cars. Labor economists, on the other hand, have long believed that workers who travel longer distances are compensated for these longer commutes, which implies that reductions in commuting duration from improved transit should depress wages. We expect there to be additional effects as land and housing prices shift in response to the transit system and people and firms relocate to take advantage of the access provided by the new system. Because these individual effects sometimes work in opposing directions, the net effect of improved mass transit on labor market outcomes is likely to be city-specific and can only be settled empirically.

---

[1] Eligible commute trips include those that are not walk trips and are less than 25 kilometers.
[2] Walking distance is characterized here as those transportation analysis zones (TAZs) that have some portion of their area within 1000 meters of a station.

The current paper provides a rare empirical assessment of whether a major investment in urban transit infrastructure actually improved employment outcomes across a city, and if so, for whom. In this paper, we use a repeated cross-section labor market dataset to assess whether access to the new BRT system affects the incomes of those who live in station area neighborhoods. Our results indicate that the opening of the TransMilenio system was associated with increased income for those living near – but not immediately adjacent to – trunk line stations. This relationship is strongest in the lower and middle-income range.

The paper proceeds as follows. The following section situates the paper by reviewing the related theoretical and empirical literature. Section three provides an overview of the data sources used in our analysis. Section four presents our empirical approach and results, and Section five concludes.

2. Theory and previous evidence

This study explores how incomes change within a neighborhood in response to an exogenous change in public transportation infrastructure. In the existing literature, there are a variety of hypotheses about these relationships (see Gibbons and Machin 2006 for a good summary of this topic). In this section, we summarize these hypotheses and a selection from the empirical literature that has aimed to find support for them.

The primary effect of a new public transport station in a neighborhood is to change the cost of travel for residents of the neighborhood who use public transportation, and to make the neighborhood more accessible to all users of public transportation in the city. This cost of travel can be decomposed into two components: the money cost (e.g., the bus fare) and the time cost. In the case of the TransMilenio, travel times decreased without a substantial increase in fares for those whose origins and destinations were served by the new system (Sandoval and Hidalgo 2004). This primary effect can cause secondary effects, including changes in local land (and therefore housing) prices, changes in physical development patterns near stations, and changes in wages and employment for people with improved job access. Some of these effects are likely to change the type of household that is most attracted the neighborhood, causing in- and out-migration and changing the neighborhood composition. Not surprisingly, the direction of the net effect of all these changes on neighborhood incomes and employment levels is not immediately obvious.

2.1. Housing Prices and Public Transport

There is a sizable literature that focuses on estimating the effect of public transportation infrastructure on land and housing prices in the vicinity of stations. If public transportation is considered to be a positive amenity for neighborhood residents, then land prices should rise when a new station is built to reflect the improved location. A rise in land value would logically lead to increases in housing prices and/or increases in housing unit density, though existing empirical studies have looked almost exclusively at housing prices rather than density. In some cases, especially in the immediate vicinity of a new station, public transport may be seen as a disamenity for the area as the increased noise, crowding, and crime outweigh the positive effect of reduced public transport travel times. If this is the case, then land prices would drop, with a likely associated drop in housing prices (and possibly a decrease in housing unit density as housing is replaced by commercial space). Overall, findings have been that where there is a statistically significant effect of public transport on housing prices, it is positive (see RICS 2002 and Salon and Shewmake 2010 for useful reviews of this literature in developed and developing cities, respectively).

There have been a number of studies specifically looking at the effect on housing prices of the public transport system that is our focus – the TransMilenio in Bogotá. In Bogotá, we expect that the amenity effect would outweigh the disamenity effect in most neighborhoods because of the high percentage of households that do not own cars

(73%) and who are transit commuters (72%) (*Encuesta de Movilidad* 2005). The findings in existing research are positive, but mixed in terms of their statistical significance. Rodriguez and Targa (2004) and Rodriguez and Mojica (2008) find that some property values have risen substantially as the system has been built. They can statistically attribute these increases to proximity to TransMilenio's dedicated busways, but there is substantial variation in the effect by neighborhood. Munoz-Raskin (2010) has similarly mixed findings, and shows that proximity to feeder routes has a larger positive economic impact than proximity to a trunk line.

2.2. Wages and Commute Costs

Also related to our research question, there is a separate literature within urban labor economics that focuses on estimating the effect of commute cost on wages. There are two theoretical constructs that lead to the same basic conclusion: workers with longer commutes are compensated with higher wages. This implies that an improvement in public transportation infrastructure that reduces commuting costs should actually cause incomes to *decrease, ceteris paribus*. Because commute length is only reduced for those who use public transport, these theories predict that this negative effect on incomes would be largest for low- and middle-income workers.

The first of these theoretical constructs is based on an extension of the classical Alonso-Muth-Mills monocentric city model that allows employers to locate somewhere other than the center of the city. The main result of this model is that employers locating farther from the city center will pay lower wages for the same work, and a spatial "wage gradient" will emerge. If a large percentage of an area's job opportunities are located in or near the center of a city, then this result can be seen as a spatial income gradient for households as well. Households living farther from the city center will earn higher incomes, on average. White (1999) gives a thorough survey of the theory and evidence for wage gradients, focusing on the employers' locations. Manning (2003) looks at the household income gradient, finding that long commutes are only partially compensated for by wages.

The second of these constructs is a spatial extension of the “efficiency wages” model. This model starts from the assumptions that workers will shirk if they can get away with it, that shirking is costly to employers, and that monitoring is imperfect. The resulting equilibrium features high, downwardly rigid wages; firms find that paying more than the going rate for workers of a given type reduces shirking because it makes the threat of firing more costly. Zenou and Smith (1995) extend this model to reflect the effects of space. The result of this extension is that workers produce only if they are compensated for their commutes, or what the authors refer to as “space costs”. They go on to show that, in equilibrium, both wages and the unemployment level increase in space costs. In one germane empirical application of this model, Ross and Zenou (2008) use US data to show that efficiency wages describe outcomes for blue-collar workers, but not white collar workers. For these lower-income workers, longer commutes correlate with higher levels of unemployment and higher wages.

2.3. Job Search Costs, Job-Worker Matching, and Public Transport

The third strand of literature that is relevant for our research develops models of the impact of improved transportation on job searches and job-worker matching. Urban search theory posits that search frictions are the fundamental source of unemployment in the city. Search models depart from traditional models of the labor market – which assumed that a worker could instantly and costlessly choose to work as many hours as she chose at the market wage – by recognizing that finding a job takes time, and that both workers and firms must evaluate matches of varying quality. Zenou 2011 develops a search-matching model that reflects conditions in Bogota rather well – he explicitly considers the role of rural-to-urban migration and evaluates the effect of an investment in the public transportation system, relative to comparably sized government investments that subsidize firms' entry costs or that restrict migration. If the impact of commuting costs on job creation is low, he shows that reducing per-kilometer commuting costs decreases wages (due to lower spatial compensation), *reduces* land rents throughout the city (due to lower accessibility costs to the job center), and increases urban employment. While the model developed in Zenou 2011 is particularly relevant for Bogota, it is a special case of a general class of models of urban search-matching that enjoys substantial empirical support (Petrongolo and Pissarides 2001).

The urban search model reviewed above misses a potentially important dynamic: improvements in public transportation infrastructure may improve matching between jobs and workers by lowering search costs and increasing search radii. These better-matched workers are more productive, which should lead to higher wages. The impact of commute time on the matching technology has been explored theoretically (see Zenou 2009, Chapter 7) in the context of the spatial mismatch hypothesis. The spatial mismatch hypothesis holds that spatial arrangements in American cities contribute to poor labor market outcomes for disadvantaged populations. In particular, the hypothesis posits that low-skilled, non-white workers live far away from jobs, and that this spatial arrangement contributes to poor labor market outcomes for these individuals. Zenou 2009 develops a model where public transportation for (poor) blacks is subsidized. He concludes that the net effect of the policy on blacks' wages depends on the relative magnitudes of two effects: the increase in wages due to improved matching and the decrease in wages due to decreased spatial compensation.

2.4. Incomes and Public Transport

In a study that is similar in important respects to our work, Glaeser et al (2008) explore the relationship between household incomes and proximity to public transport in US cities. To estimate this relationship, they construct a panel of urban census tracts, some of which experienced new public transportation access during the period of the panel. They find that tracts that were "treated" with new transit access experienced small decreases in median income levels and small increases in poverty rates. This result was robust to a range of fixed effect specifications. A 2010 study of 42 recently-opened public transport rail stations in 12 metropolitan areas in the US finds that their effect on neighborhood incomes is more mixed (Pollack et al. 2010). Median neighborhood incomes rose faster than those in the surrounding metropolitan area in approximately two-thirds of the newly transit-accessible neighborhoods.

In sum, existing theoretical and empirical results give contradictory guidance regarding the expected effect of an exogenous change in public transportation costs on wages, employment, and location decisions. All the models reviewed imply that some households are likely to relocate when commuting times change. Some of them also suggest mechanisms by which change in commute times could affect the incomes of existing residents. The unresolved state of the literature points to the importance of empirical research. Changes in public transportation infrastructure unleash a complex and sometimes contradictory set of forces, the net effect of which can only be determined empirically. In the remainder of this paper, we explore the relationship between incomes and proximity to a newly-opened public transport station in Bogotá, a city where most residents use public transportation regularly.

3. Data

We use data from two main sources. The first is a labor market survey that covers approximately 2000 households in Bogotá each quarter (the *Encuesta Continua de Hogares*, or ECH, which is carried out by DANE, the Colombian national statistics office). The survey is ongoing, but we have obtained data from the years 2000-2005, inclusive. Data prior to 2000 are not comparable to post 2000 ECH rounds, and post-2005 data were not available to our research team (Arango, Garcia and Posada 2008). The observations in this data are geographically identified at the level of the *manzana*, a small spatial unit roughly equivalent to a census block in the United States. In Bogotá, many *manzanas* are approximately the size of a city block. This labor market survey provides us with basic socioeconomic data as well as information about the changes in employment status and income over time.

The ECH survey data is collected continuously, repeatedly drawing random samples of households from a subset of the *manzanas* in the country. This means that the survey design is a repeated cross-section rather than a true panel: there are many repeated observations for each *manzana* in the dataset, but there are seldom repeated observations for individual households.

Our subset of the ECH data is based on those employed individuals who reported their monthly wage income.[3] Table 1 shows some key summary statistics from the portion of the ECH dataset on which our analysis is based. The first column provides average values for income and four additional indicator variables for individuals in our dataset who reside more than 1500 meters from a TransMilenio station. The next two columns indicate the values of these variables *before* the opening of the TransMilenio for those living between 750 and 1500 meters from and within 750 meters of the (not yet open) station locations. The final two columns indicate the *change* in these summary statistics that occurred *after* the stations opened. As is clear from Figure 1, the TransMilenio system did not open all at once - different BRT lines opened at different times. Accordingly, Table 1 presents averages for the periods before and after the nearest station to each individual's home opened, regardless of opening date.

[3] About 56% of the 209,055 observations in the original dataset were dropped due to their unemployment status. We leave for future work the inclusion of unemployed individuals in our analysis, which requires the specification of corner-solution or selection models. The size of our final sample (58,835) also reflects adjustments due to missing values for crucial variables such as age, education level, and type of occupation. Data from the first quarter of 2000 were not considered reliable and therefore also not considered. Finally, individuals not reporting income and those in the bottom and top 0.5% of the income distribution in each quarter were not considered in our estimations.

Table 1: Selected summary statistics for the full dataset

		Pre-TM		Change After TM	
	Not Near TM	TM1	TM2	TM1	TM2
Average Income (1000s)	415	555	470	**-140**	**+110**
Percent Female HH Head	26	27	27	+1	+1
Percent >Secondary Education	26	45	35	**-12**	**+9**
Percent Low Status Job	78	60	68	**+12**	**-4**
Average HH Size	4.3	4.0	4.2	+0.1	**-0.2**
Number of observations	36,239	3,896	4,545	5,907	9,404

Notes: Boldface type indicates that the difference in this variable between pre- and post-TransMilenio is statistically significant at the 95% level.
TM1 indicates observations within 750 meters of a TransMilenio station.
TM2 indicates observations located 750-1500 meters from a TransMilenio station.

Tables A-1 and A-2 provide these summary statistics separately for observations that are near each of the TransMilenio lines.

There are two interesting points to note in Table 1. First, the statistically significant changes that occurred after stations opened are largely consistent with the neighborhood becoming poorer in the distance band that is closest to stations, and wealthier in the distance band between 750 and 1500 meters from a TransMilenio station. In this second band, there are more educated workers, fewer workers with low status jobs[1], average household size has fallen, and actual incomes have risen. As will become clear, this is entirely consistent with our regression results, and may be an indication that much of the effect we find is the result of neighborhood composition changes, likely due to household relocation.

The second data sources are GIS maps. We have a map of the TransMilenio system, including information about when each bus station and feeder route first opened, a map of all of the roads in the city, and maps of all levels of the census geography in Bogotá, including the *manzana* (block) level. Together with the labor market data, these maps allow us to calculate the distances between survey respondent homes and TransMilenio stations.

4. Empirical approach and results

In this study, we aim to answer two related research questions. First, we ask whether the opening of TransMilenio BRT stations had a positive or negative effect on the incomes of those residing near them. Second, we explore whether this effect differed depending on the income level of the local residents. We find that proximity to the new BRT stations did have a positive effect on the incomes of those living in low- and middle-

[1] We categorized job categories as low or high status in the following way. Low status jobs included driver, construction or factory worker, farmer, security or police officer, server, hotelier, office worker, clergy, and low-level sales and management. High status jobs included professional, medical worker, teacher, creative, business owner, foreman, insurance, and high-level sales and management.

class neighborhoods. Interestingly, this effect was statistically significant in areas that are near BRT stations, but not immediately adjacent to them. Here we describe our empirical approach to arrive at these results and provide a real-world interpretation of what they suggest.

For our main empirical analysis, we employ a straightforward log-linear weighted least squares (WLS) regression specification. The weights are given by the inverse of the probability of being sampled. This approach does not require averaging across people or households, and thereby allows us to take full advantage of the fact that our sample is large. Our dependent variable is the natural logarithm of individual wage income measured in constant 2000 Colombian pesos. We estimate models with two distinct sets of explanatory variables, described in detail below.

Table 2 presents the results for eight WLS regressions. The odd-numbered (or "base") regressions include only location and time-related covariates as explanatory variables. The even-numbered (or "full") regressions also include sociodemographic and job-specific information about the individuals in the sample. The location-related covariates control for the spatial heterogeneity of neighborhoods, and include the distance to a major road, the distance to the Central Business District (CBD), and dummy variables for each city district (there are 20 districts – or *localidades* – in Bogotá). Most of these variables are statistically significant. Distances to a major road and to the CBD are negatively correlated with income, consistent with our expectation that households living in more accessible locations tend to be wealthier. The time-related covariates are simply dummy variables for the year and quarter of the survey date for each observation.

To identify the impact of proximity to the TransMilenio system, we include additional explanatory variables related to each household's home location relative to TransMilenio stations. To do so, we first assign households to one of three spatial bands around each station: within 750 meters; between 750 and 1500 meters; and more than 1500 meters from a station. These bands are roughly equivalent to walking times of less than 10, between 10 and 20 minutes, and more than 20 minutes respectively. We then include four independent variables that indicate the proximity of household home locations to the TransMilenio stations both before and after stations are opened: *tm1, tm2, postm1, and postm2.* The variables *tm1* (within 750 meters), and *tm2* (between 750 and 1500 meters) are coded 1 if the home is in the corresponding distance band for at least one station and zero otherwise – *regardless of whether the station was open or not* at the time of the survey. The variables *postm1* and *postm2* use the same distance bands as above, but are coded to 1 *only if the station is open* at the time of the survey. The estimated coefficients on each of these variables should be interpreted as relative to the relationship between income and living more than 1500 meters from a station. *Postm1* and *postm2* are our main variables of interest; they isolate the relationship between income and proximity to the TransMilenio stations. Including *tm1* and *tm2* serves to control for differences between neighborhoods before the opening of the system, insuring that our results are related to system access rather than features of the neighborhoods where the stations were placed.[2]

The additional explanatory variables in the "full" regressions are the more traditional determinants of individual income, such as education level, age, and part-time worker status. Not shown in Table 2, these regressions also include 24 categories for type of job, and 23 categories for the economic sector where the person is employed. Most of these coefficients are statistically significant and appear with the expected signs.

---

[2] Note that a household could be located in *tm1* and *tm2* for one or more stations simultaneously. We implemented another set of regressions that included variables that take into account the number of stations to which the home is close to. The results are not different from the ones presented in Table 2.

Table 2: Weighted Least Squares Model Results (dependent variable: natural logarithm of income)

	Full sample		Low and Medium estratos		High estratos	
	(1)	(2)	(3)	(4)	(5)	(6)
tm1	0.0626	-0.0002	0.1254**	0.0429	-0.1187	-0.1099
	(0.0640)	(0.0282)	(0.0611)	(0.0284)	(0.0980)	(0.0772)
tm2	0.0489	0.0335	0.0649	0.0460*	-0.0328	-0.0360
	(0.0503)	(0.0278)	(0.0454)	(0.0246)	(0.1481)	(0.1424)
postm1	-0.0770	-0.0384	-0.0949	-0.0465	0.0732	0.0788
	(0.0741)	(0.0305)	(0.0703)	(0.0317)	(0.1867)	(0.1061)
postm2	0.1314**	0.0212	0.1136**	0.0265	-0.0228	-0.0953
	(0.0609)	(0.0314)	(0.0516)	(0.0278)	(0.1535)	(0.1457)
cbd dist	-0.0852***	-0.0434***	-0.0499***	-0.0240***	-0.0093	-0.0086
	(0.0077)	(0.0041)	(0.0078)	(0.0034)	(0.0117)	(0.0078)
pvial dist	-0.2601***	-0.1095**	-0.3107***	-0.1410***	0.3607	0.1130
	(0.0911)	(0.0453)	(0.0716)	(0.0361)	(0.2454)	(0.1853)
edu		0.3889***		0.3551***		0.3332***
		(0.0123)		(0.0121)		(0.0446)
age		0.0477***		0.0499***		0.0626***
		(0.0018)		(0.0018)		(0.0064)
age squared		-0.0005***		-0.0006***		-0.0006***
		(0.0000)		(0.0000)		(0.0000)
female		-0.1252***		-0.1279***		-0.0891***
		(0.0078)		(0.0079)		(0.0325)
parttime		-0.7853***		-0.7986***		-0.4910***
		(0.0116)		(0.0119)		(0.0399)
head		0.1373***		0.1383***		0.3033***
		(0.0077)		(0.0076)		(0.0333)

constant	14.3338***	13.2798***	13.6061***	12.8015***	13.4600***	12.4404***
	(0.1338)	(0.0934)	(0.1741)	(0.0982)	(0.1802)	(0.2061)
PLUS: Location and Quarterly time trend variables in all regressions. Industry and Occupation variables in regressions (2),(4) and (6)						
Observations	58835	58835	52418	52418	4487	4487
R-squared	0.1471	0.5261	0.0925	0.4878	0.1172	0.5534

Sampling weights used in estimations

Cluster-robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01

Before embarking on a discussion of our complex results regarding the relationship between proximity to TransMilenio stations and incomes, we first review our main hypotheses about this relationship and explain the rationale for including both our "base" and "full" regressions. There are two main reasons why proximity to a TransMilenio station may be correlated with higher incomes. First, these locations now have improved access to all areas of Bogotá served by the system, making it possible for residents to access employment opportunities that were previously unavailable to them. This should improve matching between employers and employees, leading to higher productivity and thereby higher incomes. Second, these locations have a new accessibility amenity, increasing the value of properties in the neighborhoods and attracting higher income households to relocate to these areas.

These two hypotheses are respectively related to the “area” and “sorting” effects in the literature on spatial wage disparities (Combes et al. 2008, Gibbons et al. 2010). In the former story, existing residents of the TransMilenio neighborhoods get different jobs and earn higher wages. In the latter story, existing residents of TransMilenio neighborhoods are displaced by higher income households for whom TransMilenio access is more valuable. To the extent that we include covariates in our regression specification to control for demographics, job type, and industry, we are likely to be partially controlling for the sorting effect on incomes. Our "full" regressions include these controls, but our "base" regressions do not. If the main effect on income is due to sorting, including these variables should reduce the estimated effect of proximity to TransMilenio stations. It is hardly surprising, then, that the estimated coefficients on our variables of interest – *postm1* and *postm2* – are smaller in our full than in our base regressions.

From regression (1), the main conclusions regarding access to transit and incomes is that compared to being more than 1500 meters from a station, location within

750 meters of a station has no effect on income, but there is a positive correlation between the second distance band and income once stations are opened. Incomes in the neighborhoods between 750 and 1500 meters from an open TransMilenio station are 13% higher than incomes in areas not served by the new system. From regression (2) we can see that the positive effect of *postm2* shrinks and becomes statistically insignificant once these individual characteristics are included as covariates. This could be an indication that higher-income people are moving to neighborhoods within the second distance band instead of incomes rising as a result of improved mobility. This finding is in line with the results in Combes et al. (2008) and Gibbons et al. (2010), where sorting of individuals with different skills are the main drivers of wage disparities across regions.

4.1. Poor versus rich

The regressions in Table 2 differ not only in the covariates explaining income but also in the subsample considered for estimation. This latter differentiation allows us to see how our results differ across income groups. Our full sample is used to estimate regressions (1) and (2), while (3) to (6) use only portions of the sample based on the *estrato* level in which the individual lives. In Bogotá, the *estrato* level is an indicator of the services available at each *manzana* and is used by the government for differentiation of tax rates and public services fees. Every *manzana* is assigned an *estrato* level from 1 to 6. *Estrato* level is roughly (but not exactly) correlated with neighborhood affluence. Importantly, even though transportation improvements took place during the period analyzed, the *estrato* level did not change for any of the *manzanas* considered in our analysis. This provided us with an adequate exogenous variable on which to base the construction of our subsamples. We grouped *estratos* into low and medium (2, 3 and 4) and high (5 and 6). Regressions (3-4) and (5-6) are respectively based on these *estrato* categories.[1]

The results observed in our full sample are also present in the low-medium *estrato* subsample regressions (3) and (4). On the other hand, no *postm* coefficients are significant in regressions (5) and (6) for individuals in high *estratos*. This suggests that to the extent that TransMilenio affected incomes, this effect was mainly felt in low and medium *estrato* areas. This is consistent with our expectations. Those living in high *estrato* neighborhoods are most likely to own and use cars for transportation, meaning that the opening of a TransMilenio station does not appreciably change the accessibility of these neighborhoods for those who live there.

4.2 Job type, education level, and employment status

In order to complement our income analysis we also estimated the probability of being a high status worker depending on the location of the individual's home. The reason we looked at the probability of being in a high status job is that if people living in the vicinity of a TransMilenio station are in different job categories before and after it opens, this can be interpreted as additional evidence that perhaps the correlation between incomes and proximity to TransMilenio is due to people moving rather than people getting higher-paying jobs in their same job categories.

Estimates from probit models are presented in Table 3. Results for our full sample indicate that the probability of being a high status worker increased after stations opened in the second band. Similar to the regressions in Table 2, the positive coefficient in the probit for the full sample seems to be driven by the low-medium *estrato* subsample for which the *postm2* coefficient is also positive.

---

[1] The number of observations in *estrato* 1 that are within the two distance bands before and after TransMilenio is very small. We discarded these data from the subsample analysis.

Table 3. Probit models (dependent variable: high status)

	Full sample	Low and Medium estratos	High estratos
tm1	0.1502*	0.2135***	0.0025
	(0.0799)	(0.0793)	(0.0863)
tm2	0.1008*	0.1129*	0.0857
	(0.0595)	(0.0590)	(0.1092)
postm1	-0.1136	-0.1226	-0.1725
	(0.0904)	(0.0923)	(0.2594)
postm2	0.1463**	0.1201*	0.0216
	(0.0712)	(0.0661)	(0.1310)
cbd dist	-.1046***	-0.0816***	-0.0078
	(0.0093)	(0.0123)	(0.0120)
pvial dist	-0.3320**	-0.4046***	-0.0825
	(0.1363)	(0.1190)	(0.2757)
constant	1.3079***	1.0585***	0.1542
	(0.1572)	(0.2521)	(0.1917)
PLUS: Location and Quarterly time trend variables in all regressions.			
Observations	58835	52418	4487
R-squared	0.0972	0.0713	0.0447

Sampling weights used in estimations
Cluster-robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01

Table 4. Probit models (dependent variable: education)

	Full sample	Low and Medium estratos	High estratos
tm1	0.1528*	0.2195***	0.0175
	(0.0845)	(0.0846)	(0.1310)
tm2	0.0928	0.1020	0.0582
	(0.0660)	(0.0653)	(0.1016)
postm1	-0.0820	-0.1120	0.1249
	(0.0964)	(0.0980)	(0.2667)
postm2	0.2444***	0.2098***	0.1645
	(0.0798)	(0.0752)	(0.1231)
cbd dist	-0.1216***	-0.0949***	-0.0083
	(0.0114)	(0.0152)	(0.0142)
pvial dist	-0.4500***	-0.5701***	0.0312
	(0.1624)	(0.1512)	(0.4145)
constant	1.9300***	1.3378***	0.4015
	(0.1915)	(0.2995)	(0.2355)
PLUS: Location and Quarterly time trend variables in all regressions.			
Observations	58835	52418	4487
R-squared	0.1329	0.1029	0.0608

Sampling weights used in estimations
Cluster-robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01

Table 4 shows the results from a model that estimates the probability of whether individuals completed some education beyond high school. For both the full sample and the low-medium *estrato* subsample, location within the second distance band after stations opened has a positive and significant effect on the probability of individuals having higher education. Assuming that educational attainment was not directly affected by access to TransMilenio in the period analyzed, this result would be indicative of people with higher education moving into the second distance band. Based on the statistical significance of models not presented here, TransMilenio did not affect the spatial distributions of gender and age.

We also estimated the probability of being employed explained by dummy variables for distance to TransMilenio, location and quarterly time-trend. Observations on the unemployed labor force were appended to our dataset. These regressions (results not presented) indicate that none of the variables representing distance to opened and unopened TransMilenio stations were statistically significant in predicting employment.

4.3 Discussion

Examined together, our results suggest that proximity to an open TransMilenio station does appear to have a statistically significant positive effect on earnings of individuals in these areas, particularly those in low and medium *estratos*. However, when controlling for individual factors that affect income, the TransMilenio effect on income vanishes. Interestingly, the main effect appears not in the immediate vicinity of the station, but instead in the distance band between 750 and 1500 meters from a station. One plausible—but speculative—explanation for this result is as follows.

Close to new stations, increased congestion and commercial activity actually reduce the desirability of residential locations, causing some wealthier residents to move away. In addition, redevelopment might occur close to stations to replace single-family homes with smaller high-density residential apartments marketed to low- and middle-income households. Even if some existing residents are able to change to higher paying jobs due to the improved access offered by the TransMilenio, the potential for lower income households to move in will make it hard to identify this effect.

A short distance away, however, the negative externalities caused by the station are largely absent and the stations are far enough away that station area redevelopment effects are also absent. Residents would mainly experience the positive effect of improved transit access together with an enriched local commercial center near the station. These improved amenities could attract higher income households to the neighborhood, as well as providing existing residents with access to better job opportunities.

4.4. Limitations of our study

While this story is consistent with our results, we do not have the data required to gain a clear understanding of why the effect of the TransMilenio is strong only in the area that is a short distance away from the stations. A more basic challenge we face in this study is the potential for simultaneity between income and proximity to TransMilenio stations. Being near a BRT station may lead to better job access and thereby higher income and/or the amenity value of BRT stations may attract higher income households to relocate to these areas. Due to the fact that our data is a repeated cross-section rather than a true panel, we cannot establish with confidence which of these explanations represents the truth.

Panel data would solve this problem by providing information on the residence location of an individual across time, but it is not available. Alternatively, information about the years that households have resided in their current location could be used to restrict the analysis to those observations that remained in the same area before and after the transit improvement. Unfortunately, this information is, likewise, not available.

However, as should be evident from our review of the literature, identifying a clear relationship between income and proximity to transit remains a worthy contribution to the literature in this area – regardless of the causal mechanism behind that relationship.

Unfortunately, this same issue of simultaneity also leads to the potential for estimation bias that occurs whenever the dependent variable and an independent variable in a regression model are co-determined. We considered a number of alternative estimation strategies in the hope that we might identify a way to circumvent this limitation of our data. Aggregating our data geographically and using panel methods does not solve this problem because populations within each area might not be stable due to relocation. Similarly, pseudo-panel methods applied to cohort averages (Verbeek and Nijman 1992) do not circumvent the issue of the potential bi-directional causation between income and access to the BRT stations. To use techniques such as propensity score matching and instrumental variables methods, what is needed is a variable that affects location but not incomes, but such a variable does not exist. Our results, then, should be interpreted as highly suggestive rather than conclusive.

5. Conclusions

Our analysis suggests that the TransMilenio system did result in a statistically significant increase in income for households living near, but not immediately adjacent to, trunk-line stations. This finding is robust to several alternative specifications, but because we do not have true panel data, we cannot fully discern whether this observed relationship is due to household relocation decisions, to improved labor market outcomes for a constant set of households, or to a combination of these effects.

That said, the evidence that we do have suggests that moving is an important part of the story. This evidence includes both statistically significant results from our probit models of job status and education level, as well as the attenuated parameter estimates for the effect of transit access on wages in regressions that include a full set of sociodemographic controls. Additionally, a set of regressions on subsamples of our data separated by *estrato*, a variable that classifies *manzanas* by their level of public services availability, suggests that the positive effect of TransMilenio on incomes was concentrated in areas where lower and middle-class households live.

Regardless of the cause, our results stand in contrast with recent empirical and theoretical investigations into the effect of commuting cost on household wages. These studies have found that reducing commute costs actually *lowers* wages, and that proximity to public transport *lowers* incomes. If validated by other empirical studies, our finding that proximity to the TransMilenio in Bogotá is associated with *increased* household incomes could have subtantial implications for transportation policy.

Further work is needed to establish the external validity of our results. Several cities are currently planning large-scale investments in mass transit. This provides an unusual opportunity to carry out comparative studies of the effect of public transit on labor markets in different contexts, and ultimately to arrive at an empirical description of how the characteristics of urban transit systems might enhance their positive effect on labor market outcomes.

Acknowledgments: This research was supported by grants from the Earth Institute at Columbia University and from the Sustainable Transportation Center at the University of California Davis, which receives funding from the U.S. Department of Transportation and Caltrans, the California Department of Transportation, through the University Transportation Centers program.

References

Arango, L.E., Garcia, A.F. & Posada, C.E., 2008. The methodology of the Continuous Household Survey and the connection of the Colombian series urban labor market. *Desarrollo y Sociedad*, 207–248.

Cain, A., G. Darido, M.R. Baltes, P. Rodriguez, and J.C. Barrios. 2006. Applicability of Bogotá's TransMilenio BRT System to the United States. Final Report to the Federal Transit Administration, US Dept. of Transportation. Project Number FL-26-7104-01. 86 pp.

Combes, P.P., G. Duranton, and L. Gobillon. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2): 723-742

Departamento Administrativo Nacional de Estadistica (DANE). 2000-2005. *Encuesta Continua de Hogares*, data files.

Departamento Administrativo Nacional de Estadistica (DANE). 2005. *Encuesta Continua de Movilidad*, data files.

Echeverry, J.C. et al., 2005. The Economics of TransMilenio, a Mass Transit System for Bogotá. Economia, 5(2), 151-196.

Gibbons, S. & Machin, S., 2006. Transport and labour market linkages: Empirical evidence, implications for policy and scope for further research. Background paper for the Eddington report to the Department of Transport.

Gibbons, S., H. Overman, and P. Pelkonen. 2010. Wage disparities in Britain: People or place. SERC Discussion Papers 0060, Spatial Economics Research Center, LSE.

Glaeser, E. L., M. E. Kahn, and J. Rappaport. 2008. Why do the poor live in cities? The role of public transportation, *Journal of Urban Economics*, 63(1): 1-24.

Heres, D.R., D. Jack, and D. Salon. Do public transit investments promote urban economic development? Evidence from Bogotá, Colombia. Presentation at the 2010 Kuhmo-Nectar Conference on Transport Economics, Valencia, Spain.

Institute for Transportation and Development Policy (ITDP). 2003. Bus rapid transit spreads to Africa and Asia. *Sustainable Transport*. 15: 4-11.
Manning (2003)

Munoz-Raskin, R.C. 2010. Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia. *Transport Policy*, 17(2): 72-84.

Petrongolo, B. and C. Pissarides. 2001. Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, 39(2): 390–431.

Pollack, S., B. Bluestone, and A. Gartsman. 2010. *Maintaining Diversity In America's Transit-Rich Neighborhoods: Tools for Equitable Neighborhood Change*. Report for the Dukakis Center for Urban and Regional Policy at Northeastern University.

RICS Policy Unit. 2002. *Land Value and Public Transit*. Office of the Deputy Prime Minister, London, UK.

Rodriguez, D. A. and F. Targa. 2004. Value of accessibility to Bogotá's bus rapid transit system. *Transport Reviews*, 24(5): 587–610.

Rodriguez, D.A. and C.H. Mojica. 2008. Capitalization of BRT network effects into land prices. Working Paper, Lincoln Institute of Land Policy.

Rogers, C.L., 1997. Job Search and Unemployment Duration: Implications for the Spatial Mismatch Hypothesis. Journal of Urban Economics, 42(1), 109–132.

Rogerson, R., Shimer, R. & Wright, R., 2005. Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4), 959–988.

Ross, S. L. and Y. Zenou. 2008. Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory, *Regional Science and Urban Economics*, 38(5): 498-517.

Salon, D., D.R. Heres, and D. Jack. Do public transit investments promote urban economic development? Evidence from Bogotá, Colombia. Presentation at the 2009 Association of Collegiate Schools of Planning Annual Conference, Crystal City, VA.

Salon, D. and S. Shewmake. 2010. Opportunities for value capture to fund public transport: A comprehensive review of the literature with a focus on East Asia. Report prepared for the Asian Development Bank, 35 pp. Available at www.itdp.org.

Sandoval E. and D. Hidalgo. (2004), "TransMilenio: A High Capacity - Low Cost Bus Rapid Transit System Developed for Bogotá, Colombia" Proceedings of 2nd International Conference on Urban Public Transportation Systems: Ensuring Sustainability Through Mass Transit, Alexandria, Virginia April 14-18: 37-49.

TransMilenio website: www.transmilenio.gov.co

Verbeek, M. and T. Nijman. 1992. Can cohort data be treated as genuine panel data? *Empirical Economics*, 17(1): 9-23.

White, M.J. 1999. Urban areas with decentralized employment: Theory and empirical work. Chapter 36 In *Handbook of Regional and Urban Economics*, Vol. 3: 1375-1412.

Zenou, Y. and T.E. Smith. 1995. Efficiency wages, involuntary unemployment and urban spatial structure, *Regional Science and Urban Economics*, 25, 547-573.

Zenou, Y. 2009. *Urban Labor Economics*. Cambridge: Cambridge University Press.

Zenou, Y. 2011. Search, migration, and urban land use. The case of transportation policies, *Journal of Development Economics*, 96: 174-187.

Appendix: Summary statistics stratified by TransMilenio line

Table A1: Selected pre-TransMilenio summary statistics by line

	Line A			Line B			Line D			Line F			Line H		
	Not Near Line A	TM1	TM2	Not Near Line B	TM1	TM2	Not Near Line D	TM1	TM2	Not Near Line F	TM1	TM2	Not Near Line H	TM1	TM2
Average Income (1000s)	440	534	542	431	534	722	457	370	518	456	404	454	473	320	316
Percent Female HH Head	26	36	33	26	33	28	26	30	24	26	21	28	26	27	25
Percent >Secondary Education	30	55	43	29	40	45	31	30	45	31	30	33	33	19	16
Percent Low Status Job	74	52	63	75	56	61	73	76	60	74	72	67	72	79	83
Average HH Size	4.2	3.6	3.7	4.2	3.9	4.1	4.2	4.1	4.4	4.2	4.4	4.1	4.1	4.1	4.4
Number of observations	55,404	77	380	56,245	82	694	54,827	183	336	55,674	781	1,381	52,552	799	905

TM1 indicates observations within 750 meters of a TransMilenio station.
TM2 indicates observations located 750-1500 meters from a TransMilenio station.

Table A2: Change in selected summary statistics by line between pre- and post-TransMilenio

	Line A		Line B		Line D		Line F		Line H	
	TM1	TM2	TM1	TM2	TM1	TM2	TM1	TM2	TM1	TM2
Average Income (1000s)	-22	**+168**	**+223**	**+174**	+21	-63	+27	+36	-27	+26
Percent Female HH Head	-5	-3	-9	+2	-4	+3	**+8**	+1	+2	0
Percent >Secondary Education	-5	**+9**	**+22**	**+21**	0	-8	**+12**	**+16**	0	**+8**
Percent Low Status Job	+10	+7	-6	**-14**	-1	**+10**	-2	-3	+3	-2
Average HH Size	-0.3	-0.2	-0.4	**-0.4**	+0.1	**-0.3**	-0.2	**-0.3**	+0.1	-0.1
Number of observations	740	2,659	427	1812	1,605	2,309	453	971	2,730	2,274

Boldface type indicates that the difference in this variable between pre- and post-TransMilenio is statistically significant at the 95% level.
TM1 indicates observations within 750 meters of a TransMilenio station.
TM2 indicates observations located 750-1500 meters from a TransMilenio station.