

ABSTRACT

Title of dissertation: THE IMPACT OF THE WASHINGTON METRO ON DEVELOPMENT PATTERNS

Katja Pauliina Vinha, Doctor of Philosophy, 2005

Dissertation directed by: Professor Nancy Bockstael
Department of Agricultural and Resource Economics

Professor Maureen Cropper
Department of Economics

It is a tenet of urban planning that transportation networks help shape the spatial configuration of cities. In the case of heavy rail systems, a common belief is that building a subway system will promote employment and population density, thereby discouraging urban sprawl and its negative consequences. This dissertation examines the impact of the Washington Metro rail system in 1990 and 2000 on the distribution of employment and population in two counties in the Washington, DC metropolitan area—Montgomery County and Prince Georges County. It asks whether employment and residential construction increased more rapidly near Metro rail stations than in other parts of the metropolitan area. It also examines the impact of the Metro on the socio-demographic composition of population near Metro stations.

Evaluating the impact of the Metro system on employment and population density is complicated by the fact that stations along the Metro line may be located in areas of

high population and/or employment density to begin with, or in areas with significant amounts of developable land available. To deal with this issue I use a propensity score matching estimator. The technique is an improvement over the traditional methods of evaluation as it acknowledges the endogeneity of the location of Metro stations. Furthermore, matching estimators relax the functional form assumptions of OLS estimators.

The research finds statistically significant impacts on employment and overall development density from proximity to a Metro station and does not find consistent impacts on population or dwelling unit densities. However, for Prince George's County a negative impact on the percentage of the population belonging to a minority is found. The results also suggest that impacts on development are greater closer to the station than farther away and that they are greater the longer the stations have been in operation.

THE IMPACT OF THE WASHINGTON METRO ON DEVELOPMENT PATTERNS

by

Katja Pauliina Vinha

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Committee:

Professor Nancy Bockstael, Co-chair
Professor Maureen Cropper, Co-chair
Professor Antonio Bento
Professor Lori Lynch
Professor Mahlon Straszheim

© Copyright by
Katja Pauliina Vinha
2005

DEDICATION

To Elvi Kulmakorpi
and
in memory of
Jenny Vinha

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the help, support and encouragement of many people. First I would like to thank my dissertation committee members. My two main advisors were exceptional. Nancy Bockstael and Maureen Cropper provided detailed feedback on my work, nudged me in the right direction when I was lost and were willing to submit to long teleconference calls to discuss my progress. I thank you for patiently waiting as I tried to find my topic. Antonio Bento always showed enthusiasm for what I was doing and gave feedback on the work that significantly improved the dissertation. The comments of both Marlon Straszheim and Lori Lynch helped to better focus the research.

Also, I would like to specially thank Jeffrey Smith whose comments on propensity score matching estimators greatly clarified my thinking. I am also indebted to Rodney Green (Howard University), Wayne Koemple (MNCPPC-Montgomery County), Philip Taylor (MNCPPC-Prince George's County) and Robert Stedman (WMATA) for sharing data that was crucial in the development of the research.

Many people provided me a home away from home while commuting from Colombia to Washington. The hospitality of the Galindo-Carvajal family, of the Gonzalez-Krivonos family and of Maria Lucia Guerra made my trips not only productive but fun.

This endeavor would not have been possible without the support of my family. My parents, Irja and Matti, provided me with a stimulating environment to question and supported me through the long dissertation process. They always believed that I would finish. My husband, Felipe Barrera, was my number one fan throughout and gave me the emotional support that kept me going during the many low points. He was willing to read and comment on my work (and sometimes explain what was it that I was doing) even if my responses were not always the best. For sure, without him this dissertation would not exist. And finally to Bebe-bebe who gave me an ultimatum to finish and literally kicked me to get done. I love you.

TABLE OF CONTENTS

List of Tables.....	vii
List of Figures.....	viii
Chapter 1: Introduction and Literature Review.....	1
1.1 Statement of the problem.....	4
1.2 Literature review – Land use and transportation interactions.....	7
1.2.1 Hedonic price studies.....	8
1.2.2 Intensity and distribution of land use.....	10
1.3 Contribution of the Dissertation	14
Chapter 2: Propensity Score Matching Estimators.....	16
2.1 Binary propensity score matching estimator	17
2.2 Multiple treatment matching propensity score estimator	32
Chapter 3: Application of the Matching Estimator.....	39
3.1 Endogeneity of station location decisions	39
3.2 Considerations in the implementation of propensity score matching.....	45
3.2.1 Differences in the two Counties.....	46
3.2.2 Differences among station areas	50
3.2.3 Temporal Impacts	61
3.2.4 Spatial Impacts.....	63
3.3 Outcome measures.....	66
3.3.1 Employment.....	67
3.3.2 Population and socio-economic characteristics	67
3.3.3 Overall Density	69
3.4 Explanatory variables in the propensity score estimation	70
3.4.1 Location of the TAZ	71
3.4.2 Population and socio-economic characteristics	72
3.4.3 Employment.....	74
3.4.4 Land use and zoning	75
3.4.5 Past growth.....	76
3.4.6 Accessibility.....	76
Chapter 4: Case Study - Montgomery County.....	78
4.1 Initial Conditions in Montgomery County.....	78
4.2 Propensity scores	90
4.3 Construction of the counterfactual.....	98
4.4 Impacts of Metro stations in Montgomery County.....	107
4.4.1 Impacts in 2000.....	108
4.4.1.1 One-mile treatment common support sample	108

4.4.1.2	One-mile treatment thick support sample	112
4.4.1.3	Half-mile treatment thick support sample.....	118
4.4.1.4	One-mile multiple treatment analysis	121
4.4.2	Impacts in 1990.....	124
4.4.3	Western branch of Red Line	126
4.5	Conclusions.....	129
 Chapter 5: Case Study - Prince George's County.....		131
5.1	Initial Conditions in Prince George's County.....	131
5.2	Propensity Scores.....	143
5.3	Construction of the counterfactual.....	150
5.4	Impacts of Metro stations in Prince George's County	156
5.5	Conclusions.....	165
 Chapter 6: Conclusions.....		166
 Appendices.....		174
 List of References.....		179

LIST OF TABLES

Table 2.1:	Algorithm for calculating multiple treatment impacts	38
Table 3.1:	Characteristics of the station areas and the rest of the county	43
Table 3.2:	Boardings by station and development potential	53
Table 3.3a:	Share of Metrorail commuting by station, Montgomery	56
Table 3.3b:	Share of Metrorail commuting by station, Prince George’s	57
Table 3.4:	Characteristics of Metro station areas (in operation in 2000)	59
Table 3.5:	Descriptive statistics of initial conditions	71
Table 4.1:	Descriptive statistics of initial conditions around Montgomery County Metro stations by year opened for 2000 TAZs	89
Table 4.2:	Propensity score calculation for Montgomery County	91
Table 4.3:	Propensity score distribution for Montgomery County	95
Table 4.4:	Optimal bandwidth for Epanechnikov kernel	100
Table 4.5:	Comparison of initial conditions in unweighted and weighted samples, 2000 TAZ	102
Table 4.6:	Comparison of initial conditions in unweighted and weighted samples, 1990 TAZ	103
Table 4.7:	Average treatment impacts for one mile radius in 2000 for Montgomery County	109
Table 4.8:	Average treatment impacts for thick support TAZs in 2000 for Montgomery County	114
Table 4.9:	Average treatment impacts for thick support TAZs in 2000 for Montgomery County when controls are at least 1.5 miles from a station..	117
Table 4.10:	Treatment impacts in 2000 with varying doses of treatment for Montgomery County	123
Table 4.11:	Average treatment impacts for thick support TAZs in 1990 for Montgomery County	125
Table 4.12:	Average treatment impacts for one mile radius in 2000 and 1990 for Western Branch stations in Montgomery County	128
Table 5.1:	Descriptive statistics of initial conditions around Prince George’s County metro stations by year opened for 2000 TAZs	142
Table 5.2:	Propensity score calculation for Prince George’s County	144
Table 5.3:	Propensity score distribution for Prince George’s County	149
Table 5.4:	Optimal bandwidth for Epanechnikov kernel	151
Table 5.5:	Comparison of initial conditions in unweighted and weighted samples, 2000 TAZ	153
Table 5.6:	Average treatment impacts for one mile radius in 2000 for Prince George’s County	158
Table 5.7:	Average treatment impacts for thick support TAZs in 2000 for Prince George’s County	162
Table 5.8:	Average treatment impacts for thick support TAZs in 2000 for Prince George’s County when controls are at least 1.5 miles from the station ...	164
Table 6.1:	Comparison of impacts from OLS and propensity score matching	173

LIST OF FIGURES AND MAPS

Figure 4.1:	Distribution of propensity scores, Montgomery County	96
Figure 5.1:	Distribution of propensity scores, Prince George’s County	149
Map 3.1:	Metro station locations in Maryland	52
Map 3.2:	Land use in Montgomery and Prince George’s counties	54
Map 4.1:	Population density in Montgomery County, 1970.....	80
Map 4.2:	Percentage of minorities in Montgomery County, 1970	81
Map 4.3:	Average household income in Montgomery County, 1970	82
Map 4.4:	Employment density in Montgomery County, 1972	85
Map 4.5:	Zoned land use in Montgomery County, 1961	86
Map 4.6:	Land use in Montgomery County, 1973	87
Map 4.7:	Propensity scores for Montgomery County	94
Map 4.8:	Thick support TAZs in Montgomery County	97
Map 4.9:	Weights for controls in Montgomery County, 2000	106
Map 5.1:	Population density in Prince George’s County, 1970	132
Map 5.2:	Percentage of minorities in Prince George’s County, 1970	134
Map 5.3:	Average household income in Prince George’s County, 1970	135
Map 5.4:	Employment density in Prince George’s County, 1972	138
Map 5.5:	Zoned land use in Prince George’s County, 1961	139
Map 5.6:	Land use in Prince George’s County, 1973	140
Map 5.7:	Propensity scores for Prince George’s County	147
Map 5.8:	Thick support TAZs in Prince George’s County	148
Map 5.9:	Weights for controls in Prince George’s County, 2000	155

Chapter 1

Introduction and Literature Review

It is a tenet of urban planning that transportation networks help shape the spatial configuration of cities. In the case of heavy rail systems, a common belief is that building a subway system will promote employment and population density, thereby discouraging urban sprawl and its negative consequences. This dissertation examines the impact of the Washington Metro rail system in 1990 and 2000 on the distribution of employment and population in two counties in the Washington, DC metropolitan area—Montgomery County and Prince Georges County. It asks whether employment and residential construction increased more rapidly near Metro rail stations than in other parts of the metropolitan area. It also examines the impact of the Metro on the socio-demographic composition of population near Metro stations.

Potential densifying impacts from a subway system may arise if either employers or employees prefer to locate near a Metro station. If being close to a Metro station reduces the time or money costs of commuting, improves accessibility, improves the neighborhood air quality, or increases the potential market area for firms, one would expect a concentration of development in these areas at the expense of development in other areas. When the stations are built in areas with higher potential for future development, development may be increase relative to other areas. That is, there is a re-distribution of the development and hence a slow-down in sprawled, or decentralized, development. This redistribution may impact different income groups differently if, for

example, the improved accessibility to jobs is not constant for all types of employment and if the type of housing built favors certain income groups over others.

It is also possible that such densification does not occur or that there is a densification of only employment or population. It is possible that households prefer low-density housing, in which case there may not be a demand for higher-density housing units. Also, if the time savings are sufficiently great households may choose to locate farther out and drive to the station to take advantage of the Metro service. In addition, there may be negative impacts from being located closer to a Metro station, such as higher crime rates, increased traffic volumes from commuters who drive to the station, or higher noise levels, which could reduce development densities or attract lower income housing.¹

Evaluating the impact of the Metro system on employment and population density is complicated by the fact that stations along the Metro line may be located in areas of high population and/or employment density to begin with, or in areas with significant amounts of developable land available. As is true of project evaluations in general, one cannot simply compare employment and population densities in areas near to and far from Metro stations after a Metro is built because station locations are endogenous. To deal with this issue I use a propensity score matching estimator. The technique is an improvement over the traditional methods of evaluation as it acknowledges the endogeneity of the location of Metro stations. Furthermore, matching estimators relax

¹ For example, Ihlanfeldt (2003) finds evidence of increases in crime rates near stations where the income of the surrounding neighborhoods is low in Atlanta.

the functional form assumptions. Specifically, I estimate an equation to predict the probability that a Metro station will be built near (or in) a transportation analysis zone (TAZ), as a function of variables describing TAZs in the early 1970s, before the Washington Metro was built. This equation is used to identify control TAZs—ones that did not have Metro stations built nearby but which are observationally similar to TAZs that are near a Metro station (treatment TAZs). I then compare levels of employment and population density and other variables between treatment and the identified control TAZs in 1990 and 2000.

It should be emphasized that the goal of the dissertation is to examine the impact of a Metro rail system on the neighborhoods where Metro stations were located. It is not possible to construct a counterfactual of not having a rail transit system in place in order to determine the impact of the transit system on the regional growth. Therefore, the results encompass both relocation within the region and any additional growth that the system itself generates.

In order to examine the question, this dissertation is organized as follows. The rest of Chapter 1 sets out the research question of the dissertation and reviews the existing studies on the interaction of land use and transportation. It concludes by outlining the contribution of this dissertation to the existing literature.

Chapter 2 discusses the foundations of the empirical methodology that I use. It describes the theory behind propensity score matching estimation—with only one, and

with several different, treatments. Chapter 3 gives an overview of the actual Metrorail network in Montgomery and Prince George's counties as well as a description of the planning process in the two counties. It also describes the empirical approach adopted to determine the impacts of the Metro system on land use patterns. Furthermore, the Chapter describes how the different variables in the empirical analysis are constructed and the sources of the data. Chapter 4 presents the empirical results for Montgomery County. Chapter 5 presents the results for Prince George's County. Chapter 6 concludes with the principal findings of the research.

1.1 Statement of the problem

The effectiveness of an urban transit system in changing urban shape depends on how it alters the incentives that households and employers face. A heavy rail system provides an alternative method of transport by joining various parts of the metropolitan area via a network of lines and stations, potentially reducing the time and money costs of transport. If the heavy rail system affects the transport costs of distinct areas differently, then the relative costs of displacement within the metropolitan area change and the choice of housing and employment locations may also change. Certain areas may become relatively more attractive, because of transport savings, and the demand for these areas will increase also increasing the land rent.² This increase in rents possibly changes the optimal development density (for example, Amin and Capozza, 1993). If the increase in possible rents is greater than the demolition and construction costs, then we would expect

² Other factors affecting development patterns are planning and public policies as well as geographical attributes. That is, policies such as zoning and the construction of highways influence how a metropolitan area develops. Zoning limits the type of land use in the short run. In the long run zoning ordinances may change reflecting the current needs. In this dissertation these aspects are subsumed into the locational attributes.

higher densities around the subway stations. However, it is also possible that the potential disamenities from transit station affect the optimal decisions. It is an empirical question whether or not densification occurs and to what degree.

Furthermore, if the land rents increase it is possible that lower income population no longer can afford to live in areas with close proximity to a subway station. Thus, it is not clear that the responses to a heavy rail system on location decisions are identical across different socio-economic groups. For example, Gabriel and Rosenthal (1989) find evidence that socio-economic conditions affect differently the location choices of black and white household in the Washington, DC metropolitan area. Furthermore, in a more recent study using microdata Bayer, McMillan and Rueben (2004) find that the differences in socio-economic characteristics of various racial groups explain to a large degree the observed locational separation of different racial groups. If there are differences among different racial groups in terms of their willingness or capability of paying for locational amenities, then they may be affected differently by a heavy rail network.

The residential and commercial land markets compete for the use of space. In some cases zoning ordinances preclude the use of land for different purposes, such that only residential or commercial uses are allowed. In these cases one would possibly expect to see densification of one type of development but not the other. However both Montgomery and Prince George's counties allow for mixed-use zoning and which in

some cases require at least two different uses within an area. In these areas it is possible that both uses intensify.

It is the objective of this dissertation to examine whether or not such densification has occurred due to the construction of a heavy rail network. Specifically the dissertation asks:

1. How has the distribution of population (as measured by population density or dwelling unit density) been affected by the construction of a heavy rail transit network?
2. How has the distribution of employment (as measured by employment density) been affected by the construction of a heavy rail transit network?
3. How has overall development density (as measured by the sum of the employment and dwelling unit densities) been affected by the construction of a heavy rail transit network?

The answers to the above questions, in the context of the Washington Metrorail's impacts in Montgomery and Prince George's counties Maryland, give some indication as to whether public transit infrastructure investments along with any complementary land use policies alter incentives sufficiently to shape the spatial configuration of a metropolitan area. . Furthermore, the dissertation will also preliminarily look at the impact of the Metro on the income (and racial makeup) of the counties.

One difficulty in answering these questions is that the locations of Metrorail lines and stations are endogenous. If the placement of Metrorail stations depends on the location decision of household and firms—as it surely does—this must be taken into account in evaluating the impact of public transportation networks on urban form.

Before discussing how I will answer the questions posed above, I review the literature on the interaction between land use and transport. I focus on studies that examine the impact of public transit on land values, and on population and employment densities. In examining these studies it is important to ask whether the study has adequately controlled for the endogeneity of the transit network.

1.2 Literature review – Land use and transportation interactions

There are two strands of literature that examine the interactions between land-use or urban spatial structure and transportation infrastructure. The first looks at the impact of existing land-use (population density, employment location, and type of neighborhood) on transport decisions, and especially on commute mode choice.³ The second looks at the impact of transit systems on land use. It attempts to test the hypothesis, outlined in the previous section, that transportation infrastructure affects the location decisions of firms

³ This first question has received far more attention in the academic literature than the second. In general, in the first set of models, in general, the location of people and employment is held constant and household decisions regarding transportation are analyzed. The general hypothesis tested is that dense areas, with a mixture of land uses, and with access to public transportation networks promote the use of non-automobile forms of travel by changing the relative costs of traveling by private transportation and public transportation. That is, people living in these areas are expected to demand fewer cars and fewer private vehicle miles if the alternative of using public transportation is economically attractive. The evidence has been mixed as to whether or not urban spatial structure affects transit usage and travel behavior (for example, Baum-Snow and Kahn, 2000; Bento, Cropper, Mobarak and Vinha, 2005; Boarnet and Crane, 2001; Cameron, Kenworthy and Lyons, 2003; Cervero, 2002; Kain, 1964; Train, 1980). Badoe and Miller (2000) provide a review of work published in the transportation literature up to late 1990s.

and households. These studies focus on public transit systems, and examine their impact on land values or analyze their effect on development intensities or growth around station areas.

1.2.1 Hedonic price studies

The hedonic price studies analyze the impact of transit on property prices. In a well functioning market, land prices should reflect the desirability of a particular area. In order to estimate the impact of a transit station on housing price the studies include either distance to the closest station or a series of dummy variables to represent distance from the station (Bowes and Ihlanfeldt, 2001; Cervero, 1994, 2004; Gatzlaff and Smith, 1993; McDonald and Osuji, 1995; McMillen and McDonald, 2004; Voith, 1993). Alternatively the studies look at the impact of a change in the distance to the closest station on the changes in housing prices (Baum-Snow and Kahn, 2000), or how housing prices evolved with the introduction of a transit station using repeat-sale housing prices (Gatzlaff and Smith, 1993; McMillen and McDonald, 2004). The hedonic price studies do not account for the endogeneity of subway station locations. That is, they assume that the station is placed randomly within the city. Some studies, however, limit the sample to only those housing units that are within a certain distance of a transit station (for example McMillen and McDonald, 2004).

The studies provide evidence of a premium from locating near a transit station (Baum-Snow and Kahn, 2000; Cervero and Landis, 1995; Gatzlaff and Smith, 1993; Grass, 1992; McMillen and McDonald, 2004), suggesting potential densification of development. The results, however, are not necessarily linear in space. Bowes and

Ihlanfeldt (2001) find a positive overall impact from rail stations but residential properties within a quarter of a mile of a station actually experience a negative price impact. They find the largest positive impact on price occurs in housing that is further away from the CBD and within 0.25 and 0.5 miles from the station. Gatzlaff and Smith (1993) find that in Dallas the price of housing in the higher-priced residential areas is higher if the housing is located within walking distance but not directly next to the station. They do not find any capitalization from the rail transit station in the lower-priced residential areas. These studies suggest that the transit stations are considered as amenities, but the impacts are not necessarily linear in space and may vary across socio-economic groups.

Besides distance, other factors also matter in determining the size of impacts. Voith (1993) finds that although the amount of capitalization depends on general economic conditions, the overall trend has been positive. During energy crises the premium for being close to a station is higher than “average” and during recessions this premium is much lower. Similarly, McMillen and McDonald (2004) find that in Chicago the positive impact of being close a transit station increases through time over a 17-year period. They do also find, however, that in the latest period the premium is lower than in the years immediately after the station has opened.⁴ Also, there is some evidence that prices respond anticipatorily to the announcement of a station (Damm, Lerman, Lerner-Lam, and Young, 1980; Knaap, Ding and Hopkins, 2001; McDonald and Osuji, 1995; McMillen and McDonald, 2004). These results underline the importance of time in analyzing the impacts from transit network.

⁴ They have seven years of data in the post-operational phase.

Studies on the impacts of transit stations on office rents are less numerous than on residential properties. They do not find strong evidence of a positive impact. Cervero (1994) finds some evidence that stations with complementary measures, in this case joint-development projects, increase office rents in Atlanta and Washington, DC. Furthermore, Cervero (2004) finds positive impacts on commercial properties along some (but not all) of the light rail lines in San Diego. On the other hand, Ryan (2005) does not find this rent premium for office or industrial properties in the San Diego area. The hedonic price studies thus suggest a possible densification impact on employment, but the evidence for this is much weaker than for a potential impact on residential densification.

1.2.2 Intensity and distribution of land use

Although the hedonic studies in general observe a premium on property (or land) prices, they cannot be used to determine whether or not actual densification has occurred and what type of development has densified. It may be that land has appreciated, but that the augment in price is not sufficient to induce higher density development, especially if it requires the demolition of existing properties. There are several studies that have attempted to look at the impact of stations not on prices but on the actual land use patterns.

The results from these studies are mixed, some finding positive impacts on employment (Cervero, 1994; Cervero and Landis, 1993, 1997; Green and James, 1993) and others not (Bollinger and Ihlanfeldt, 1997; McDonald and McMillen, 2000). Studies looking at the impact on residential properties or on population are far fewer, but as

inconclusive (Cervero and Landis, 1997; Bollinger and Ihlanfeldt, 1997). There are three key aspects that need to be considered in these evaluation studies – the unit of observation used, the length of time the network has been functioning, and most importantly the methodology used to determine the impacts. If the study does not carefully establish each of these aspects, the assessed impacts may be erroneous.

The first consideration is the aggregation level at which the studies are conducted. If the units are large and encompass very heterogeneous areas, they may not reflect well the diversity within the unit. This may lead to very imprecise estimates. A wide range of observational areas is used in the literature. The studies tend to use Census tracts (Bollinger and Ihlanfeldt, 1997, 2003; Cervero and Landis, 1997), aggregated traffic analysis zones (Green and James, 1993), or rings (of varying radii) or impact areas around the station areas determined by planners (Cervero and Landis 1993, 1997; Metropolitan Washington Council of Governments, 1992; Moon, 1990) as the unit of observation. Census tracts are relatively big geographical areas. An average tract in a metropolitan area is six square miles. The rings vary from 0.2 square miles to 113 square miles and the aggregate traffic analysis zones for the DC region range from 0.1 square miles to 20 square miles. McDonald and McMillen (2000) use quarter sections, which are equivalent to 0.25 square miles.

The second important factor is how long the system has been operating. Given that the effects may take a long time to be realized, the short run impacts may be zero but long run impacts may be quite different. Most of the studies reviewed acknowledge this

fact and use relatively long evaluation periods. In these studies, the post-opening time period ranges from four years (some of the stations analyzed in Green and James, 1993) to twenty years (Bollinger and Ihlanfeldt, 2003; Cervero and Landis, 1997). In most studies the stations have been operating for a decade when the analyses are carried out (Bollinger and Ihlanfeldt, 1997; Cervero and Landis, 1993; the oldest stations analyzed in Green and James, 1993; Metropolitan Washington Council of Governments, 1992).

The most important aspect in evaluating the impacts is the methodology used. Given that planners locate transit stations depending on the existing land use patterns and distributions, it is important to choose non-transit areas that are similar to the transit areas in order to make meaningful comparisons. For example, studies that describe land use around station areas (Metropolitan Washington Council of Governments, 1992; Moon 1990) without references to non-station areas do not provide any possibilities to compare station areas with the rest of the urban area.

Several studies attempt to in some way identify non-station areas with characteristics similar to the station areas (Cervero and Landis, 1993, 1997; Green and James, 1993), but in general the methods of determining the control areas (or even the station areas) seem rather *ad hoc*. For example, Cervero and Landis (1993) use five station areas (two in Atlanta and three in Washington, DC) and compare these to non-station areas that were similar in development density and mix prior to the opening of the station, distance and accessibility to regional centers, and in regulatory policies. It is not clear if any statistical analyses were carried out to determine the comparison area for each

station area. The authors calculate a 12-year average of several development indicators (for example, the average annual new square feet of office space, and the average square feet of office space per parcel) for each station area and compare that with the 12-year average in the chosen control area. In their later study, Cervero and Landis (1997) again choose five station areas along the San Francisco transit system and compare the averages to four areas of similar size but centered on important freeway interchanges that is within a 2.5 miles of a rail station. It is not clear statistically how similar the two areas are. For example, it appears from graphs presented that the freeway sample had much lower residential and non-residential densities prior to the opening of the urban rail system. Green and James (1993) compare total employment around Washington DC station areas to the rest of region not covered by the transit system, and rail corridors to corridors along busy non-limited access highways. The comparison with highway corridors is an improvement over the comparison with the rest of the metropolitan area, which can be very different from station areas. However, the authors do not explicitly show that the comparison areas are similar and it is not clear whether any statistical tests were carried out to confirm similarity.

The studies by McDonald and McMillen (2000) and Bollinger and Ihlanfeldt (1997, 2003) use regression analyses to determine impacts of rail transit stations on the probability of development, growth in population and employment and the change in the share of regional employment, respectively. These studies better control for additional characteristics thought to influence urban development. They do not, however, consider the characteristics of the areas prior to the opening of the transit systems since all use

post-transit data only. That is, the fact that station areas have in general very different land use from an average area within the metropolitan region is not considered in the analyses.

1.3 Contribution of the Dissertation

This dissertation adds to the existing literature on the impacts of transit stations on development patterns by using an alternative methodology that explicitly incorporates the endogeneity of station location decision into the analyses. Although prior analyses control for some initial factors and make attempts to choose control groups that are comparable, they fail to simultaneously control for a vector of factors that both determine the station location as well as the future development potential and they fail to statistically test for the similarity between the two groups of areas. This omission potentially leads to the comparison of areas that were very different even before any mass transit system was built. If this is the case, any observed impacts may reflect the pre-existing differences and not true impacts. In order to overcome this weakness, the dissertation uses matching techniques. First I explicitly model, using a discrete choice model, the Metro station location decision based on information from the early 1970s, prior to the opening of the first stations in the network. Then, using a kernel matching estimator,⁵ I calculate for the station areas the counterfactual of not being located next to a Metro station, based on the outcomes in areas with similar characteristics in 1970 but where no Metro station opened. The database describing initial conditions is built from various georeferenced data sources, and the measures themselves are constructed using geographical information system software.

⁵ I use the Epanechnikov kernel, discussed in Chapter 2.

In addition, the dissertation looks at various aspects of development. The majority of existing studies concentrate on the impact of transit stations on employment. In addition to the impact on employment, I consider impacts on population, dwelling units, and socio-economic characteristics of the population as well as impacts on the overall development density. That is, the dissertation examines the two sets of decision makers that may have been affected by the public transit system. Furthermore, given the possibility that the transit system affects different socio-economic groups differently, analyses are carried out to determine such impacts. These impacts are estimated in two Maryland counties that differ significantly in initial conditions, as well as in perceptions of the benefits of Metrorail.

In order to explore the temporal development of impacts, the dissertation analyzes land use patterns at two points in time, 1990 and 2000. By 2000, the first stations in the system had been operating for twenty-two years and the average station for fifteen years. The geographic unit of analysis used in the dissertation is the traffic analysis zone (TAZs), which ranges in area from 0.005 square miles to five square miles. (Some of the information on dwelling units is aggregated up from parcel level information.) The use of smaller geographic units should guarantee greater homogeneity in the characteristics of the area. That is, there should be lower variance in the measures; thus, they should better reflect the actual conditions in each observational unit.

Chapter 2

Propensity Score Matching Estimators

This chapter describes the theory behind propensity score matching, the approach that I use to select a set of control locations (TAZs) to compare with the TAZs where Metro stations were located.⁶ Nonparametric matching estimators are frequently applied in evaluation studies. The general idea of the methodology is to determine the impact of treatment on the treated using information from non-treated observations to build a counterfactual of not having received the treatment. I discuss the methodology for both the binary treatment case as well as for the multiple treatment case. Both are used in estimating the impacts of transit stations.

There are several reasons why traditional parametric regression analysis may not be suited for analyzing the impacts of endogenous policies.⁷ Berhman, Cheng and Todd (2004) discuss these reasons. First, the true relationship between explanatory variables and the outcome variable may be very nonlinear. Since nonparametric estimation methods do not make any assumptions on the functional form (as do parametric analyses) it is not necessary to know the exact relationship between the explanatory and the outcome variables. Second, it is possible that the participants in a particular program are quite different from the average non-participant. Using methods such as matching these differences are reduced since the control group individuals are re-weighted to better

⁶ TAZs are land areas defined by planning agencies for transport planning. I use them as the unit of observation in the empirical application.

⁷ Strictly speaking, propensity score matching is a quasi-parametric approach. Propensity scores used in constructing control groups are estimated parametrically, but treatment effects are nonparametrically determined.

match the treatment group. Third, in traditional regression analysis there may be problems of nonoverlapping support. It is possible that only treatment observations are found over certain ranges of x , and only control observations over other ranges. Traditional parametric regression analyses extrapolate the results to these regions where there are no observations. Non-parametric methods restrict the analysis to only those areas that are similar.

In the case of evaluating the impact of the stations on urban development the above-mentioned advantages of the non-parametric methodology are important. First, the exact functional relationship between density and the various neighborhood characteristics is not known. Second, the decision as to where to locate a Metro station is based on the characteristics of the location and thus the non-station areas, in general, differ in their attributes from station areas. Parametric estimation methods may attribute some differences in the impact measures to the underlying differences and not to the actual “treatment” received. Also, it is likely that there are some non-station areas that are sufficiently different from any in the station area groups that no information on a counterfactual exists. In this case parametric methods will extrapolate the results to these regions of non-overlapping support.

2.1 Binary propensity score matching estimator

One objective of program evaluation is to calculate the mean impact of the program on those treated. If the treatment condition is denoted by $T=1$ for those who received the treatment and $T=0$ otherwise, and the impact variable of interest is denoted

by y^1 if treatment was received and y^0 if not, then the objective is to estimate the equation $E(y^1|T=1) - E(y^0|T=1)$, the difference in the outcome with and without the treatment for the treated group. Unfortunately the second term is not observable, since for an individual in the treatment no outcome without the treatment exists. Thus, the challenge is to be able to say something about the unobserved counterfactual for those who have been part of the program.

When the treatment is assigned *randomly*, then it can be assumed that the covariates and unobservables do not differ in any systematic way between the treated and non-treated groups. That is, they come from the same distribution. In this case, to estimate the average treatment effect on the treated, φ , of the outcome variable, y , one can compare the after treatment outcome levels of the two groups. The average treatment impact in a randomized experiment can be calculated as:

$$\varphi = E(y^1 | T=1) - E(y^0 | T=0) \tag{2.1}$$

where the assumption is that $E(y^0 | T=1) = E(y^0 | T=0)$. That is, those in the treatment group would have had, on average, the same outcome level as the control group participants had they been assigned to the control group.

In the case of a non-randomized program, such as the building of a subway system, the treatment and control groups may vary in a systematic way, and it no longer can be assumed that $E(y^0 | T=1) = E(y^0 | T=0)$. Therefore, the treatment outcome measure for the non-participant group is not a valid counterfactual for the treatment group without treatment. As a specific example, if the location of subway stations is

based on some characteristics of the area, such as population and employment densities, then one would expect outcome measures in the treatment areas to be quite different from those in an average control area, even without the treatment. In this case, the outcome levels of the average non-treated areas are not good proxies for unobserved outcomes of the treatment group.

When there are no experimental data available, when assignment to the treatment group is non-random, and the treatment status is determined by some set of covariates, x , then an alternative mechanism needs to be employed to determine the treatment impact. One such mechanism is to establish a control group that is similar in x to the treatment group. The set of x ought to capture both the variables that affect the treatment decision, as well as those that influence the outcome measure.⁸ The average treatment effect is based on the difference in the average outcomes of the individuals in the treatment group and this “matched” control group with similar set of x .

Matching on the covariates guarantees that the two groups have similar distributions of covariates and a treatment impact mimics that of a randomized experiment. Formally, the treatment impact is captured by

$$\varphi = E(y^1 | x, T=1) - E(y^0 | x, T=0) \quad (2.2)$$

⁸ The covariates do not need to include variables that are strictly from the pre-treatment period. That is, if the objective is to analyze the impact of a job-training program on wages or unemployment rates, it is not necessary to have information on the wages or unemployment status prior to entering the program. It is assumed that by matching on factors such as formal education, age, etc. that determine wages and unemployment status these pre-treatment conditions are also captured. However, the measures included in x that may influence the outcome measure should not have been affected by the treatment. For example it would not be possible to use current population density in the set of covariates that explain current employment density given that population density may have been affected by the treatment (proximity to a subway station).

where the outcomes are conditioned on the covariates that determine treatment participation.

The above approach works only when (i) outcomes, conditional on the set of covariates, are independent of the group to which the individual belongs; and (ii) there is no covariate that unequivocally decides the treatment assignment. Mathematically, these conditions of strong ignorability⁹ can be represented as:

$$(y^0, y^1) \perp\!\!\!\perp T / x \quad \text{and} \quad 0 < Pr(T=1) / x < 1. \quad (2.3)$$

When the above conditions apply, the control group can be used to infer information about the treatment group. If there are any unobservables that influence the treatment decision and the first condition of strong ignorability does not hold, then the control group does not provide the necessary counterfactual information. The second condition rules out the possibility that any particular condition or characteristic unequivocally determines inclusion in or exclusion from the treatment.¹⁰

Rosenbaum and Rubin (1983) show that it is not necessary to match individuals based on the vector of observable characteristics, x , per se; matching on balancing scores, such as the propensity score, $b(x)$, is sufficient. The propensity score is, in effect, the conditional probability of being assigned to the treatment group given the individual's covariates. In Theorem 3, the authors demonstrate that when treatment assignment is strongly ignorable in x then it is also strongly ignorable in $b(x)$. That is if

⁹ Strong ignorability is the same as conditional independence, unconfoundedness, or selection on observables.

¹⁰ For example, it cannot be the case that all areas within x miles from the CBD are within a Metro station treatment zone and no areas farther than x miles are outside station treatment zones.

$$(y^0, y^1) \perp\!\!\!\perp T / x \quad \text{and} \quad 0 < Pr(T=1) / x < 1$$

then also

$$(y^0, y^1) \perp\!\!\!\perp T / b(x) \quad \text{and} \quad 0 < Pr(T=1) / b(x) < 1. \quad (2.4)$$

The above theorem greatly aids in the assignment of individuals into the control group since a univariate score can be used instead of a vector of individual covariates (or subclassification of the observations based on the covariates). Therefore, it is not necessary to match the observations based on multiple dimensions but only on a “summary” measure.

Rosenbaum and Rubin (1983) further show that if the treatment assignment is strongly ignorable, the average treatment effect can be obtained by comparing the treatment and matched control groups solely conditioned on the propensity score. Therefore, the treatment impact, \hat{E} , is given by:

$$\hat{E} = E\{y^1 / b(x), T = 1\} - E\{y^0 / b(x), T = 0\} = E\{y^1 - y^0 / b(x)\}. \quad (2.5)$$

The average treatment effect is the average outcome level of those in the treatment group minus the average outcome level of those in the control group after conditioning on the propensity score. The methodology, besides determining the appropriate control group to use and reducing the bias in the treatment impacts, is also desirable because it allows for the control of covariates when the sample size is small (Rosenbaum and Rubin, 1983).

The impact, however, is valid only for the observations within the common support—that is, the range of propensity scores for which there are both control and treatment observations. For example, if there are no observations with high propensity

scores in the control group, then those observations with high propensity scores are outside of the region of support. Common support, CS , is defined as the set of propensity scores for which the distributions of $T=1$ and $T=0$ have positive values, such that $CS = pdf(T = 1) > 0 \cap pdf(T = 0) > 0$. That is, the common support is the range of propensity scores for which there is a positive probability of observing both treatment and control observations. It is possible that there is no exact match for a treatment observation's propensity score. As long as within a pre-specified interval of propensity scores (i.e. within 0.05 points) there is a control observation then the two observations are said to be within the same support.

In practice, the first step is to estimate a binary choice model (logit or probit) where the dependent variable is whether or not the observation is in the treatment group and the covariates include all the variables that influence the treatment condition as well as those that may affect the impact measures. These probabilities, $\hat{P}(x)$, are then used to construct the counterfactual of no treatment for the treated based on the non-treated individuals. There are several ways to construct the counterfactual, or several methods of matching the observations. These include counterfactuals based on one control observation per treatment observation, as well as counterfactuals based on some weighted average of several control observations.

In choosing the matching algorithm, the first decision is to determine the number of control observations. On the one hand, choosing only one control observation per

treatment observation¹¹ reduces the bias that is introduced when the matched pairs are less similar in their probability of receiving treatment. On the other hand, with a greater number of comparison observations the precision of the estimates, or the magnitude of the standard errors, is better. As often the case in empirical work, the trade-off is between unbiasedness and precision.

After determining the number of observations, it is necessary to define the matching estimator, or the manner in which the counterfactual is determined for each treatment observation. The generic matching estimator for observation i in the treatment group is given by

$$E(y^0 | \hat{P}(x_i)) = \sum_{j=1}^{N_0} W(\hat{P}(x_i), \hat{P}(x_j)) y_j^0 \quad (2.6)$$

where $W(\cdot)$ determines the weight of each control observation j in the counterfactual for observation i . The various matching algorithms differ in the weights they place on the control observations to build the counterfactual.

If only one control is used per treatment observation, then the logical match for each treatment observation is the control observation with the closest propensity score, or nearest neighbor matching. In this case a weight of one is given to the control observation with the closest propensity score. That is, the treatment impact is given by

$$\frac{1}{N_1} \sum_{i=1}^{N_1} (y_i^1 - y_j^0)$$

where N_1 is the number of treatment observations, y_i^1 is the outcome for

¹¹ The control observation in a pairwise-matching will be the observation with the closest propensity score to the treatment group observation.

treatment group observation i , and y_j^0 is the outcome for the control group observation j which has the closest propensity score to observation i . The nearest neighbor to observation i is defined as observation j such that

$$\left\| \hat{P}_i(x) - \hat{P}_j(x) \right\| \leq \left\| \hat{P}_i(x) - \hat{P}_k(x) \right\| \quad \forall \quad k \in I_o \neq j \text{ where } I_o \text{ is the set of all possible}$$

control observations. For nearest neighbor matching, it is also possible to set a maximum value, d , (often called a caliper) for the difference, such that

$$\left\| \hat{P}_i(x) - \hat{P}_j(x) \right\| \leq d \text{ in order to limit the differences between treatment and control}$$

observations. A caliper can also serve as a measure for observations to be within a common support. In this case, it is possible that not all treatment observations have a control observation within this maximum difference and that particular treatment observations will thus be dropped from the analysis.¹² As noted by Smith and Todd (2004) there is no way of determining, *a priori*, an acceptable size for d .

With nearest neighbor matching, one also needs to determine whether or not to match with replacement. When matching with replacement each control observation can serve as the counterfactual for more than one treatment observation. Dehejia and Wahba (2002) show that without replacement (and without imposing a caliper) the later matched pairs can differ considerably in their propensity scores. This is especially the case when there are relatively few possible controls for some range of propensity scores. Allowing replacement, the number of “better” matches increases. However, the variance of the

¹² That is, observations are not used since they do not fulfill the common support condition.

estimator increases given that less control group information is used and it possible that several control group observations are relied upon very heavily.

When multiple controls are assigned to a given treatment observation, then it is necessary to determine how to weight the control group observations to construct the counterfactual. Adapting the notation of Heckman, Ichimura and Todd (1997, 1998), the general form to calculate the average treatment impact, $\hat{M}(T)$, can be given as:

$$\hat{M}(T) = \sum_{i \in I_1 \subset CS} \omega(i) [y_i^1 - \sum_{j \in I_0} W(i, j) y_j^0] \quad (2.7)$$

where y_i^1 is the outcome with treatment for observation i , y_j^0 is the outcome for the control observation j , and $W(i, j)$ is the weight that appears in equation (2.6). $W(i, j)$ is the weight given to observation j in the control group when comparing with observation i in the treatment group, such that $\sum_{j \in I_0} W(i, j) = 1$. That is, for each treatment observation, the weights of the controls used sum to one. I_0 and I_1 are the sets of observations in the control group and the treatment group, respectively. Only those treatment observations within the common support are used.¹³ Finally, $\omega(i)$ is the weight of each treatment observation, i , in the construction of the average treatment impact. In general $\omega(i)$ is $1/N_1$, such that each treatment observation is weighted equally in the average treatment impact.

¹³ Certain matching estimators impose the common support condition “automatically.” In other cases it needs to be explicitly defined and thus the set of observations for which the weights are determined may not include all the treatment observations. An example of the first is the kernel matching estimator and of the second the local linear matching estimator. These will be discussed in detail later in the chapter.

The different matching algorithms differ in the way that the W matrix is determined. The simpler algorithms include N -neighbor matching and radial matching. In the first, the counterfactual outcome is made up of the average of the N control group observations closest in their propensity score to the treatment observation.¹⁴ The average can be a simple average of the control group observations or an average weighted by the distance of the control group observation from the treatment observation. In radial matching an average of all the control observations with a propensity score within a certain distance, d , from the propensity score of the treatment observation is calculated. That is, the number of control observations used for each treatment observation may differ. Again, it is possible to use a weighted average instead of weighting all observations equally.

Additionally, when multiple controls are used other, more complex, matching algorithms are possible. Heckman, Ichimura and Todd (1997, 1998) propose two alternative estimators – kernel matching and local linear matching estimators – that build the counterfactual using additional information from the control group observations.

In a kernel estimator the matrix W is determined by a kernel function, $K(\cdot)$.¹⁵ Following the general notation of Smith and Todd (2004), $W(i,j)$ in this case is given by

¹⁴ The formula is a generalized formula for the matching estimator. For example, for the case of 10-neighbor matching algorithm with simple weights, the $W(i,j)$ matrix is such that for row i , the matrix has a value of $1/10$ in the columns for the ten control observations, j , with the closest propensity score to treatment observation i , and 0 otherwise.

¹⁵ In essence the kernel function, $K(\cdot)$, is a histogram but instead of determining the frequency of observations in non-overlapping intervals, the kernel estimator estimates the density using *overlapping* intervals. Kernel functions used are symmetric and $\int K(z)dz = 1$.

$$W(i, j) = \frac{K\left(\frac{\hat{P}_j - \hat{P}_i}{h}\right)}{\sum_{k \in I_0} K\left(\frac{\hat{P}_k - \hat{P}_i}{h}\right)} \text{ if } |z| < \bar{Z} \text{ and } 0 \text{ otherwise, where } z = \frac{\hat{P}_k - \hat{P}_i}{h} \quad (2.8)$$

where h is the bandwidth of the kernel, and \hat{P}_i and \hat{P}_j are the probabilities of receiving treatment for treatment observation i and control observation j , respectively, and \bar{Z} is some upper limit for a kernel value. This upper limit depends on the kernel used. There are several choices for the kernel function. They differ in the way they assign weight to observations depending on the distance of the two probabilities. For example, the rectangular kernel, which gives the same weight to all control observations (within a particular bandwidth), is

$$K(z) = 0.5 \text{ if } |z| < 1 \text{ and } 0 \text{ otherwise, where } z = \frac{\hat{P}_k - \hat{P}_i}{h}.$$

The Epanechnikov kernel, which gives more weight to control observations with similar propensity scores, is

$$K(z) = 0.75(1 - 0.2z^2)/\sqrt{5} \text{ if } |z| < \sqrt{5} \text{ and } 0 \text{ otherwise}$$

where $z = \frac{\hat{P}_k - \hat{P}_i}{h}$.

In general the choice of the kernel has been shown to have little impact on the estimated weight matrix; the choice of the bandwidth, however, does typically impact the weights (DiNardo & Tobias, 2001).

The bandwidth, h , in the kernel functions determines the interval over which positive weights are given to the control observations. A kernel with a small bandwidth will use only control observations with very similar propensity scores to that of the treatment observation. A kernel with a larger bandwidth gives weight to less similar observations.¹⁶ A sufficiently small bandwidth may not find any matches for the treatment observations. A sufficiently large bandwidth will give weight to all of the control observations such that the weight vector for a particular treatment observation will take on the shape of the kernel function.¹⁷ Given the above property of the kernel estimator, it also limits the analysis to only those observations within a common support. That is, only observations with a probability of existing in both the treatment and the control groups, given the distribution of probabilities, are included.

Given that the weight matrix (and therefore also the estimated treatment impacts) is in general sensitive to the choice of bandwidth, it is important to objectively determine the bandwidth. This can be done in several ways. The easiest way is to visually inspect the data and determine which bandwidth gives a good fit. However, one would like to determine more objectively, and in an automated manner, a good bandwidth (Härdle, 1990; Pagan and Ullah, 1999). The procedures to objectively determine the optimal bandwidth take as the basis the minimization of some global error.

¹⁶ For the same difference in the propensity scores, $\hat{P}_j - \hat{P}_i$, a larger h will decrease the numerator quotient, z , of W such that more of the control observations will fulfill the $K(\cdot)$ rule.

¹⁷ For example, if the rectangular kernel is used with a sufficiently large bandwidth, then all the control observations are used to calculate the counterfactual for each of the treatment observations, and the weight matrix would be a matrix of $1/N_0$, where N_0 is the number of control observations.

The method that has been used in the evaluation literature is that of cross-validation, or the leave-one-out method (Black and Smith, 2004; Frölich, 2004a, 2004b). The objective of cross validation is to minimize the mean squared error when estimating the outcome measure y_j based on the information from the rest of the observations y_k such that $k \neq j$. That is, the mean squared error, $\frac{1}{N_0} \sum_{j \in I_0} (y_j - \bar{y}_{(j)})^2$, is calculated for various bandwidths, where y_j is the outcome for observation j and $\bar{y}_{(j)}$ is the predicted outcome using the kernel estimator when observation j is not part of the sample. Given that the outcome without treatment only exists for the non-treated group, the measure is based on the non-treated sample. Efron and Gong (1983) summarize the methodology as consisting of the following:

- (a) deleting the points x_i from the data set one at a time; (b) recalculating the prediction rule on the basis of the remaining $n-1$ points; (c) seeing how well the recalculated rule predicts the deleted point; and (d) averaging these predictions of all n deletions of an x_i . (pg. 37)

Of the different bandwidths tested the one that minimizes the mean squared error is chosen as the “optimal” bandwidth.

The other matching estimator proposed by Heckman, Ichimura and Todd (1997, 1998) is the local linear estimator. Adapting the notation in Smith and Todd (2004) the estimator is given by:

$$W(i, j) = \frac{K_{ij} \sum_{k \in I_0} K_{ik} (\hat{P}_k - \hat{P}_i)^2 - [K_{ij} (\hat{P}_k - \hat{P}_i)] \left[\sum_{k \in I_0} K_{ik} (\hat{P}_k - \hat{P}_i) \right]}{\sum_{j \in I_0} K_{ij} \sum_{k \in I_0} K_{ij} (\hat{P}_k - \hat{P}_i)^2 - \left(\sum_{k \in I_0} K_{ik} (\hat{P}_k - \hat{P}_i) \right)^2}$$

where $K_{ij} = K((P_j - P_i)/h)$. Again, any kernel function can be used.

Asymptotically all of the matching estimators will converge since in asymptotically large datasets the matches will be perfect. However, in finite samples there are differences. There are several studies that have compared the various matching estimators. The first set uses randomized experiments where $E(y^0 | T=1) = E(y^0 | T=0)$ and compares the impacts obtained with those derived from various matching algorithms on another dataset with non-participants that were not part of the experiment. Using this methodology, Dehejia and Wahba (2002) do not find any significant differences between nearest neighbor matching and radial matching. Smith and Todd (2004) also compare different matching estimators and similarly do not find any consistent results as to the superiority between nearest neighbor matching and local linear matching with reasonable bandwidths. Based on the asymptotic properties of various estimators, Heckman, Ichimura and Todd (1997) advocate the use of local linear weights given that the estimator converges faster than kernel estimators. Frölich (2004a) using Monte Carlo studies finds, however, that ridge matching¹⁸ and kernel matching are in general superior to pair-wise matching, and that local linear matching, multiple-neighbor estimators generally perform the poorest.¹⁹ He finds that the local linear matching estimator does not perform as well as the other estimators, even if it asymptotically converges faster,

¹⁸ A weighted average of the local linear regression estimator and the Nadaraya Watson estimator.

¹⁹ Furthermore Frölich finds that the weighting estimator is sensitive to trimming and states that there currently is no way to determine the optimal trimming level. Trimming is one method of imposing the common support condition, by excluding from the analysis the tails of the probability distributions of propensity scores.

when there are regions with low density of propensity scores. He finds that when the ratio of control observations to treatment observations is large, kernel matching is a good option.

When matching is done using a propensity score measure it is also necessary to determine whether or not the resulting non-treated sample is similar in the observables to the treated sample. That is, whether or not the two samples are balanced in the observables after the appropriate matching algorithm has been applied to obtain the counterfactuals for each treatment observation. The common support condition guarantees that only observations within the range of positive probabilities for both treatment and control groups are included. Balancing tests check via the use of t-tests that the means of the covariates, x , are statistically similar in the two groups (after weighting the control group observations by the weights used to construct the counterfactual). If the two samples are not similar then additional higher order terms, such as squares of the covariates used, or interaction terms of the covariates need to be included in the construction of propensity scores until the two samples are similar (Dehejia and Wahba, 2002).

In order to obtain a confidence interval on the estimated treatment impact, bootstrapping methods are used. The standard errors are calculated by resampling the data with replacement and recalculating the treatment impact using the chosen estimator, N_B number of times. Each of the N_B samples is (potentially) different since a particular

treatment observation may appear more than once. The distribution of the N_B different average treatment impacts are used to calculate the standard error or confidence intervals.

There are three options for determining the interval. If the underlying distribution is symmetric then either the standard error of the normal distribution or the percentile based confidence interval can be used. Ordering the treatment impacts, $\hat{\theta}_i$, from the lowest to highest, the percentile based confidence interval uses the $\hat{\theta}_{N_B \cdot x/100}$ and $\hat{\theta}_{N_B \cdot (1-x)/100}$ treatment impacts as the limits for a $(100-2x)\%$ confidence interval. When the underlying distribution is asymmetric then the bias-corrected bootstrap confidence intervals yield more accurate coverage probabilities (Efron and Tibshirani, 1998). In the bias-adjusted confidence intervals the percentile based confidence limits are adjusted by a factor taking into account the proportion of times in the true estimated impact using the full sample, $\hat{\theta}$, is greater than the bootstrapped replication (Efron and Tibshirani, 1998). In effect, the confidence interval is adjusted for the difference in the median and mean impact values.²⁰

2.2 Multiple treatment matching propensity score estimator

In some cases, the treatment is not a binary condition; there may be varying doses of treatment or a set of different treatment options. Joffe and Rosenbaum (1999), Imbens (2000) and Lechner (1999, 2002) expand the analysis the use of propensity score matching estimators when there are multiple mutually exclusive treatments.

²⁰ When there is no bias, that is, 50 percent of the replications are below the true estimated impact, the bias corrected and the percentile confidence intervals are the same.

In the multiple treatment case, it is necessary to determine for the M possible treatments the M theoretically possible outcomes, Y^1, Y^2, \dots, Y^M for each individual. Again, only one of the possible outcomes is realized for each individual and the other outcomes are “missing.” The challenge is to be able to determine the counterfactual for all of those treatments that the individual did not experience.²¹

Imbens (2000) weakens the initial conditions imposed by Joffe and Rosenbaum (1999) for obtaining the average treatment impact in the multiple treatment case. He shows that it is not necessary for the treatment type to be independent of all the potential outcomes. The average treatment impacts can be estimated if there is only pairwise independence. This weaker condition (*weak unconfoundedness*) requires that the treatment type t is independent of the outcome, Y^t , when subjected to treatment t conditional on the covariates. Using the notation of Imbens (2000), if $D_i(t)$ is an indicator for each individual i such that:

$$\begin{aligned} D_i(t) &= 1 & \text{if } T_i = t \\ D_i(t) &= 0 & \text{otherwise} \end{aligned}$$

then weak unconfoundedness can be expressed as

$$D(t) \perp Y^t \mid x \quad \forall t.$$

²¹ Lechner (1999) identifies three different average impacts that can be obtained. Namely, the expected average treatment effect of being in treatment t relative to treatment s for:

- (1) a randomly chosen individual from the whole population, $\gamma_0^{t,s} = E(y^t) - E(y^s)$,
- (2) a randomly chosen individual who received either treatment t or s ,
 $\alpha_0^{t,s} = E(y^t \mid T = t, s) - E(y^s \mid T = t, s)$, and
- (3) a randomly chosen individual who was in treatment t , $\theta_0^{t,s} = E(y^t \mid T = t) - E(y^s \mid T = t)$.

The outcome Y^t is independent of whether or not treatment t is applied rather than of the treatment level per se.

Furthermore, Imbens (2000) shows that, as in the binary case, the propensity score can be used to condition the outcomes instead of the vector of observables, x . When the treatments are weakly unconfounded, then the average treatment effects are equal whether conditioning on the covariates or on the propensity score. Theorem 1 of Imbens (2000) states that

$$(i) \beta(t, r) \equiv E\{Y^t \mid r(t, x) = r\} = E\{Y \mid T = t, r(T, x) = r\}$$

$$(ii) E\{Y^t\} = E\{\beta(t, r(t, x))\}$$

where $r(t, x)$ is the generalized propensity score. That is, the conditional expectation of the impact evaluated at a particular treatment level, $\beta(t, r)$, is equal to the average treatment impact, $E\{Y^t\}$.

Given that there is a propensity score associated with each of the M treatments, more than one propensity score needs to be determined for each individual. That is, each individual needs to be evaluated for her propensity to receive each of the different treatments. Lechner (2002) describes two different ways – a structural approach and a reduced approach – of calculating the propensity scores. The first estimates the probabilities using a multinomial, or ordered, discrete choice model. The predicted probabilities from the model are used to calculate the conditional probabilities

$$\hat{P}_i^{s/ls}(x) = \frac{\hat{P}_i^s(x)}{\hat{P}_i^s(x) + \hat{P}_i^t(x)} \quad (2.9)$$

where $\hat{P}_i^s(x)$ is the predicted probability of receiving treatment s given the vector of characteristics x . The conditional probabilities are required since the comparisons to be made are between two different groups and not all groups at the same time.

In the reduced approach separate binary choice equations are estimated for each of the possible $M*(M-1)/2$ pairs²² of treatments in order to obtain $\hat{P}_i^{s/ts}(x)$. That is, only observations that received either treatment t or s are included in the calculation of the conditional probability. Lechner (2002) advocates the use of this second approach on two counts. First, in the ordered multinomial probit “if one choice equation is misspecified all conditional probabilities could be misspecified” (pg. 210), given that the probabilities are all evaluated at the same time. Second, it is easier to estimate binary models than ordered models. Lechner (2002) finds that the estimated conditional probabilities are highly correlated across the two approaches and thus the treatment impacts are very similar regardless of which approach is used to estimate the propensity scores.²³

For the multiple-treatment case the common support set is in general determined by the minima of the maximum and the maxima of the minimum participation probabilities for the various treatment options (Frölich, Heshmati & Lechner, 2004). Equations 2.10 and 2.11 give the common support conditions for the lower bound and the upper bound, respectively.

²² Where M is the number of different groups, including the no treatment group

²³ All correlation coefficients for his sample were greater than 0.98.

$$Lower\ bound = \max\left(\min\left(\hat{P}_i^{s/ts}(x)\right)\forall t, s \in T\right) \quad (2.10)$$

$$Upper\ bound = \min\left(\max\left(\hat{P}_i^{s/ts}(x)\right)\forall t, s \in T\right) \quad (2.11)$$

For example, if there are three distinct treatment groups and the lowest probability of receiving treatment C is 0.1 in among those observations belonging to treatment group A and it is 0.05 among observations in group B, and 0.01 for those in treatment group C, then all those observations with a probability of receiving treatment C less than 0.1 are dropped from the sample. The procedure is applied to all of the different treatments.

Because with multiple treatments it is necessary to match on more than one conditional probability, in general, the matching is done using a nearest neighbor algorithm. The treatment impact is given by

$$\hat{M}(T) = \frac{1}{N_t} \sum_{i \in I_t, i \in CS} [y_i^t - \sum_{j \in I_s} W(i, j) y_j^s] \quad (2.12)$$

where the $W(i, j)$ is 1 for observation j in treatment s that is the $\min(d(i, j) \forall j \in I_s)$, where $d(i, j)$ is the closeness of the two conditional probabilities $\hat{P}_k^{s/ts}(x)$ and $\hat{P}_k^{t/ts}(x)$ for $\forall k \in \{I_t, I_s\}$.

The distance metric generally used in the literature is the Mahalanobis distance.²⁴

Formally, the Mahalanobis distance $d(i, j)$, between observations i and j is defined as:

$$d(i, j) = (P_i^t - P_j^s) V^{-1} (P_i^t - P_j^s)$$

²⁴ There are not many applications of multiple treatment matching. Frölich, Heshmati and Lechner (2004), Lechner (2002) use the Mahalanobis distance as the metric to determine the nearest neighbor. Behrman, Cheng and Todd (2004) use local linear regression estimators, where the weights are given by the closeness of the observations in terms of the observable characteristics and dose.

where P_i^t is a vector of propensity scores for treatments t and s for observation i in treatment group t , P_j^s is the same vector of propensity scores for observation j in the alternative treatment group s . V is the covariance matrix based on the all the subset of observations from I_t and I_s such that,

$$V = \{(N_t - 1)V_t + (N_s - 1)V_s\} / (N_t + N_s - 2)$$

where N_k is the number of observations in treatment k , and V_k is the sample covariance of the relevant propensity scores, P , in group k , $k = t, s$ (Rubin, 1980).

As a summary, the algorithm proposed by Lechner (1999) for calculating the impact of different treatments is given in Table 2.1.

Table 2.1: Algorithm for calculating multiple treatment impacts

Step 1	Specify and estimate a multinomial choice model to obtain $[\hat{P}_N^0(X), \hat{P}_N^1(X), \dots, \hat{P}_N^M(X)]$
Step 2	<p>Estimate the expectations of the outcome variables condition on the respective balancing scores. For a given value of m and l the following steps are performed:</p> <p>a) Compute $\hat{P}_N^{lml}(X) = \frac{\hat{P}_i^l(X)}{\hat{P}_N^l(X) + \hat{P}_{Ni}^m(X)}$ or use $[\hat{P}_N^m, \hat{P}_N^l(X)]$ directly. Alternatively step 1 may be omitted and the conditional probabilities may be directly modeled (as in the binary case).</p> <p>b) Choose one observation in the subsample defined by participation in m and delete it from that pool.</p> <p>c) Find an observation in the subsample of participants in l that is as close as possible to the one chose in step a) in terms of $\hat{P}_N^{lml}(X)$ or $[\hat{P}_N^m, \hat{P}_N^l(X)]$. In the case of using $[\hat{P}_N^m, \hat{P}_N^l(X)]$ “closeness” can be based on the Mahalanobis distance. Do not remove that observation, so that it can be used again.</p> <p>d) Repeat a) and b) until no participant in m is left.</p> <p>e) Using the matched comparison group formed in c) compute the respective conditional expectation by the sample mean. Note that the same observations may appear more than once in that group.</p>
Step 3	Repeat step 2 for all combinations of m and l .
Step 4	Compute the estimate of the treatment effects using the results of step 3.
Source: Lechner, 1999.	

Chapter 3

Application of the matching estimator to evaluate the impacts of transit stations

As discussed in Chapter 1, the fundamental challenge in estimating the impacts of a heavy rail system on development patterns is solving the problem of endogenous station location. Metro stations are not located randomly and the areas in which they are located may be quite different from the characteristics of an average area in the region. Not accounting for these factors will not give the correct impacts. I adopt the matching methods described in Chapter 2 in order to estimate the development impacts of the Metrorail built to serve Washington, DC and the surrounding counties. What follows is a discussion of the main aspects of the application. First, I discuss the problem associated with the endogeneity of the location of Metro stations. Second, I address the special considerations that must be taken into account in the application of the matching estimator in this case. Finally, I describe the outcome variables of interest, as well as the construction of variables used to match the treatment and non-treated areas.

3.1 Endogeneity of station location decisions

The challenge in estimating the impact of any heavy rail transit system on the distribution of the population and employment is that transit lines and stations are not chosen randomly within a region, and the characteristics upon which the decisions are

based are not static through time.²⁵ The intentional siting precludes the simple comparison of densities in station and non-station areas. Planners decide where to build stations taking into account the costs of construction and the potential market of users as defined by the actual and projected distributions of people and employment.

The Washington DC Metrorail system was planned to alleviate growing congestion in downtown employment centers by reducing peak hour commute trips in private cars. In various proposals for the system, the Washington, DC downtown area—where more than half of regional employment was located—was always envisioned as the center of the network (Murin, 1971). Given that the aim of the system was to reduce traffic congestion by providing alternate means for employees residing in the suburbs to reach downtown jobs, areas closer to the District were more likely to receive a station given the network nature of the transit system.

A second important consideration in the siting process was the location of population. The planning staff determined that in order for the system to be successful “it must be oriented to areas of population concentration at one end and concentrated employment on the other” (Murin, 1971, pg 77). That is, a key element in the siting of suburban transit stations was the concentration of potential users. This not only meant that population densities needed to be high, but also that the population in the targeted areas would utilize the services. For this reason, no service was finally planned to the densely populated north central part of the District in order to eliminate “an

²⁵ The application here is similar to studies such as Jalan and Ravallion (2003) in which the authors determine the health impact of piped water when the location of the water pipe lines is determined by public entities.

‘uneconomic’ part of the system and provide more highly utilized and more profitable service on the rest of the system” (Murin, 1971, pg. 76). In Prince George’s County due to the resistance of the wealthier neighborhoods along the planned northern branch of the Green Line, a less profitable alignment was chosen for the branch (Hanna et al., 1994).²⁶

The third factor included in the analyses was employment density. In the case of the Washington Metro, the suburban counties in Maryland and Virginia were given veto power over the location of their stations (Murin, 1971). They could effectively choose the locations that corresponded most closely to the County’s other planning objectives. In the two Maryland counties, the Metro system was developed concurrently with the adoption of comprehensive land use guidelines. In 1962 the Maryland-National Capital Park and Planning Commission (MNCPPC) developed a general plan in which urban development would be concentrated along four different corridors (one in Montgomery County and three in Prince George’s County) leaving the areas in between for lower intensity development and open space. The plan’s aim was to guide development in an orderly manner and to decrease traffic congestion by taking advantage of existing, or planned highways and transit networks. To achieve these goals various corridor cities with employment and services (including a transit facility) were envisioned. These cores, 4 miles apart, were to be surrounded by concentric rings of high-density, medium-density and low-density residential development to offer various types of housing opportunities. The plan was updated in 1969 in Montgomery County after the location of transit stops had been decided (Montgomery County Planning Board, 1969). The corridors of the

²⁶ This alignment, for example, did not go through either downtown College Park nor was it located on the University of Maryland campus, two of the potentially high use areas.

original plan coincided with the general location of the proposed transit lines in Prince George's County. Given the general plan, the transit station locations may have been influenced not only by high population densities but also by high concentrations of employment.

Besides actual concentrations of population and employment, existing land use and zoning plans were important in the siting process. They reflected past development and future development potential. Although part of the land use characteristics are captured by population and density measures these do not capture, for example, the amount of undeveloped land available for new development. It is conceivable that the amount of undeveloped land was also a factor in assessing the future development potential of the areas chosen for corridor cities. Furthermore, densification impacts depend on the availability of land—either land for redevelopment or vacant land; however, vacant land, in general, is easier to develop.

Another variable that describes the profitability of a station is past growth trends. Areas with recent development signal current locational preferences of households or firms. Potentially, if undeveloped land exists, future growth is more likely to occur in these areas, increasing the potential future market of the system.

Another factor considered was accessibility to the Metro station, especially for terminal stations or other stations with large parking facilities. Not all of the users of the system were predicted to walk or bike to the station. Some would drive to take the Metro

to the District. Therefore, stations needed to be easily accessed from other parts of the county. Ease of accessibility also captures the future development potential of an area since it measures, together with the distance from the center, the costs of transport to the center.

All of these factors in the decision-making process yield non-random station locations. Defining station areas as those areas within a mile of a transit stop, Table 3.1 shows the average population, employment, and land use characteristics for station areas and for the remainder of the two counties. The data confirm that stations were placed in higher density, more developed parts of the county.

Table 3.1: Characteristics of station areas and rest of the county

	Montgomery		Prince George's	
	Areas within a mile of a station ¹	Rest of the county	Areas within a mile of a station ¹	Rest of the county
Population density per hectare in 1970	21.3	2.9	19.0	0.4
Employment density per hectare in 1972	1.5	0.8	0.5	0.1
Pct of area in low density residential in 1973	0.078	0.086	0.017	0.026
Pct of area in medium density residential in 1973	0.421	0.073	0.320	0.101
Pct of area in high density residential in 1973	0.054	0.011	0.080	0.016
Pct of area in commercial activity in 1973	0.153	0.013	0.136	0.025
Pct of area in industrial activity in 1973	0.001	0.017	0.004	0.003
Pct of area in institutional uses in 1973	0.072	0.017	0.034	0.026

The population density measure is based on the 1970 Census and the employment on the 1972 employment used in Green and James (1993). The land use data come from the Maryland PropertyView CD.

¹ Areas within a mile of a station are composed of those traffic analysis zones (TAZs) that have their centroid within a mile of a transit station. All the other TAZs in the county are considered as part of the rest of the county

Given the path-dependent structure of development, it is necessary to carefully choose the base year for which to construct the propensity score measure. In evaluating the impact of programs to deliver piped water (Jalan and Ravallion, 2003) or rehabilitation programs (Frölich, Heshmati and Lechner, 2004), the matching can be done using only information from the post-treatment period or using time invariant characteristics. In this case, one cannot rely on post-treatment characteristics since many of the factors involved in the decision process are also (potentially) altered by the treatment. For example, if population density was important in determining station locations and there is a positive impact on population, then using post-treatment population density would incorrectly identify control areas. This implies that it is necessary to use pre-treatment information to determine potential controls for the treatment observations.

Using data from 1979, when the first stations opened in Montgomery and Prince George's counties, may not capture these initial conditions well. It is possible that by 1979 land markets had already anticipated the arrival of the system.²⁷ If the matching is performed using information altered by anticipatory behavior then areas surrounding stations may, for example, appear denser than they would have been without the announcement. If, on the other hand, pre-treatment conditions are selected very much before the planning process terminated it is possible that the conditions do not adequately reflect the actual conditions on which decisions were based.

²⁷ Gatzlaff and Smith (1993) do not find an impact on housing prices from the announcement of a transit station in Miami; however, others such as Damm et al. (1980), Knaap, Ding and Hopkins (2001), and McDonald and Osuji (1995), have found significant anticipatory behavior in Washington DC, Portland and Chicago. Damm et al (1980) find price elasticities with respect to year of completion from -0.05 to -0.15 for the Washington DC Metro.

In order to avoid contamination by anticipatory behavior I use data from the early 1970s to capture the initial conditions. This is well before any significant anticipatory behavior would be expected. Although the system was finally approved in March 1968, some realignments were approved in February 1969 as well as others that were implemented while the system was being built (Hanna et al., 1994; Murin, 1971).²⁸ Furthermore, even for stations whose location was fixed in 1969 we would not expect any significant changes in land use, even if prices had already started to react, given that the construction of new housing or office space may take up to several years to complete.²⁹

3.2 Considerations in the implementation of propensity score matching

Before implementing propensity score matching techniques there are several differences among stations that need to be addressed to clarify the type of impacts obtained. There are two types of differences—differences between the two counties and differences among the station areas in a particular county. Furthermore, any impact will vary through time and across space. Temporal variation is more common in the traditional literature on propensity scores as the impacts of any program may intensify or weaken through time. Spatial variation, however, is not usually an issue.

²⁸ For example, the alignment for the northern part of the Green Line in Prince George's County was only decided in 1980.

²⁹ Although the average length of time spent in obtaining a building permit is variable over location and size of development, on average it has been reported to be about a year. According to statistics from the US Census Bureau, the average time from the authorization to build to the actual start of construction is approximately 0.8 months for single-family housing in the northwest over the past 28 years and approximately 1.8 months for buildings with multiple units (US Census Bureau, 2005). The actual length of construction is 7.3 months for single-family units and 12.1 months for a multi-unit structure (US Census Bureau, 2005). In total, the length of time required to obtain all approvals and finish construction, at least during the period represented in this study, was about two years.

3.2.1 Differences in the two Counties

In the two Maryland counties the Metro rail system was used in differing degrees to complement the overall planning processes. On the one hand, Montgomery County was receptive to the potential benefits from Metro stations. In a detailed study of the politics involved in the siting of the Green Line in Prince George's County, Hanna et al. (1994) conclude that in Montgomery County the "alignments and stations [were] seen as pluses, [and] built quickly" (pg 3-4). Furthermore, Green and James (1993) point out that the County "deliberately routed [its] rail lines so as to work in concert with such factors as existing concentrations and zoning initiatives to attract or generate new economic activity" (pg. 11). By 1973, sector plans (more detailed master plans) were linked with zoning plans (Citizens' Advisory Committee, 1973) and each planning area of the county had its master plan approved (MNCPPC, 1973).

On the other hand, Hanna et al. (1994) conclude that "Prince George's County officials did not see Metro as a plus but rather as a set of problems" (pp. 3-4) and favored the expansion of the road network and bus system to provide transportation opportunities to its residents. Prince George's County planned future growth in the Laurel and Bowie areas, which were not serviced by the heavy rail line (Green and James, 1993). In 1998 the County evaluated the effectiveness of its past planning policies and found a "lack of effective plan implementation, the loss of countywide perspective, and the emphasis on new development as opposed to the protection and revitalization of older, established areas" (MNCPPC, 2002, pg. 16). The position of Prince George's County reflected that of its population. In several communities there was outright rejection of a station. For

example, along the proposed Green Line, such localities as Old Town College Park, Berwyn, College Park Airport, and University Park did not want a station in their jurisdictions (Hanna et al., 1994). Other Prince George's communities, such as Cheverly, accepted the Metro station but opposed any changes in the character of the community (MNCPPC, 1977).³⁰ Overall, in Prince George's County the heavy rail network had an approval rate of 62 percent, the lowest in the suburban counties surveyed (MWCOG, 1979).

Not only were the general attitudes different in the two counties, in 1970 the socio-economic makeup of the populations differed. In 1970 average annual household income was approximately \$20,500 in Montgomery but only \$15,100 in Prince George's County.³¹ In Montgomery County about 6% of households earned less than \$5,000 and 20% earned more than \$20,000. For Prince George's County the comparable shares were 9% and 6% of the households. Furthermore, in Montgomery 47% of the population worked as professionals or executives compared to only 32% in Prince George's County. The same trend was present in educational attainment. In Montgomery 33% had at least a Bachelor's degree, whereas, in Prince George's only 17% had finished college. These factors may also influence the success of the Metro system. In an analysis of mode choice studies conducted in the late 1960s, Murin (1971) found that the probability of using public transportation decreased with increasing distance of employment from the central business district, income or occupational status.

³⁰ The notable exception was the Greenbelt community that actively sought a Metro station to their neighborhood (Hanna et al., 1994).

³¹ All the figures are based on US Bureau of Census statistics reported in the Neighborhood Change Database 1970-2000 of Geolytics, Inc.

At the time of construction of the heavy rail system, Prince George's County was experiencing a major transformation of its socio-economic characteristics. Partly because of mandatory school busing enacted in January 1973, many middle-income white families moved out while lower-income black families moved into the county. In 1970, about 5% of the population in Montgomery County belonged to a minority group, whereas, 15% of the population in Prince George's County belonged to a minority group. In 1980, the minority share had risen to 14% in Montgomery but to 41% in Prince George's. In 1990 the shares were 23% versus 57%, respectively. Even though in both counties the share of minorities was increasing, it was increasing faster in Prince George's and by 1990 minorities constituted the majority of the county population. The average family income in Montgomery County grew by 29% (in real terms) between 1969 and 1999. In Prince George's County the growth was almost half of that in Montgomery (16%). Family income in the United States grew by around 20% over the same period. In summary, the income differential between the two counties grew over the 30-year period and Prince George's was not able to keep up with average national income growth.

There were also differences in the perceived competitiveness of the two counties. In the 1970s, Prince George's County was viewed in general terms as a less competitive county to attract new businesses than Montgomery (MNCPPC, 1977). MNCPPC (1977) claims that in the 1970s Prince George's County was about 40% as competitive as other metropolitan area counties taking into account single-family housing, high-rise

apartments, commercial office space, retail sales and the quality of hotel rooms.³²

However, the two counties had a similar share of the region's employment in 1972, with Montgomery having 16% of the employment and Prince George's County 14%.³³

Furthermore, the MNCPPC study finds that overall the development process was much more costly in Prince George's County than in the other regional suburban counties, both in terms of time and money. These factors made Prince George's less attractive (with or without Metro stations) than other suburban counties in the area for developers. The study concludes that, "the general character of existing development both within and adjacent to Prince George's County and the lack of positive direction or consensus among the public sector towards future development are the major factors impeding private investment within the county" (MNCPPC, 1977, pg. 22).

In terms of meeting growth expectations, the two Maryland counties performed differently. In 1962 the MNCPPC projected that in 2000 in Montgomery County there would be 335,000 jobs and 995,000 people, whereas in Prince George's County there would be 395,000 jobs and 1,192,000 people (MNCPPC, 1962). In 2000 there were 873,321 people and 419,612 jobs in Montgomery County, while there were 801,505 people and 295,587 jobs in Prince George's County. Whereas Montgomery County was able to attract more employment than had been forecasted in 1962, Prince George's County fell short of expectations. Even in 1977 it was predicted that the population of Prince George's County would grow by 15,933 people annually in the period 1970-1985 (MNCPPC, 1977). However, in Prince George's County population growth declined

³² It is not clear from the document how the various factors are weighted to arrive at the actual competitiveness of the different counties.

³³ Based on employment numbers provided by Dr. Rodney Green and used in Green and James (1993).

drastically in the 1970s, growing by less than 7,000 people in the 1970s and by about 6,500 people annually in the 1980s.

In general the growth predictions made in the early 1960s were realized in Montgomery County but not in Prince George's County. The slowdown in development in Prince George's County has been attributed to: (1) complexities and restrictions in the development process, (2) lack of support facilities and inadequate municipal services such as sewage, (3) the blue collar image and relatively low income level of residents, (4) a high level of taxation, and (5) the availability of cheaper housing options in areas farther away from the District (Hanna et al, 1994; MNCPPC, 1977).

The above discussion points to the fact that the two Maryland counties were quite different in the pre-Metro era and differed in factors that have affected development in the post-Metro era. Thus, it is possible that any impacts of the Metro on development patterns will also differ between the two counties. First, Montgomery grew more and thus experienced more development pressure. Second, the socio-economic characteristics of the two counties differ substantially. Since it is difficult to control for all differences between the two counties, I analyze the impacts of the Metro for the counties separately.

3.2.2 Differences among station areas

Besides differences between the two counties, there is also variability among stations. First, there are several types of stations in the system. Map 3.1 depicts the location of stations in Maryland. In general, the terminal stations at the end of a line have

large parking lots and relatively greater number of boardings (Table 3.2). This is especially true for Prince George's County more so than for Montgomery County. They act as the system entry point for those living (or possibly working) farther away from the District border. The stations where large parking lots were built can be expected to behave slightly differently from stations where such structures were not built.

Map 3.1: Location of Metro stations in Montgomery and Prince George's Counties

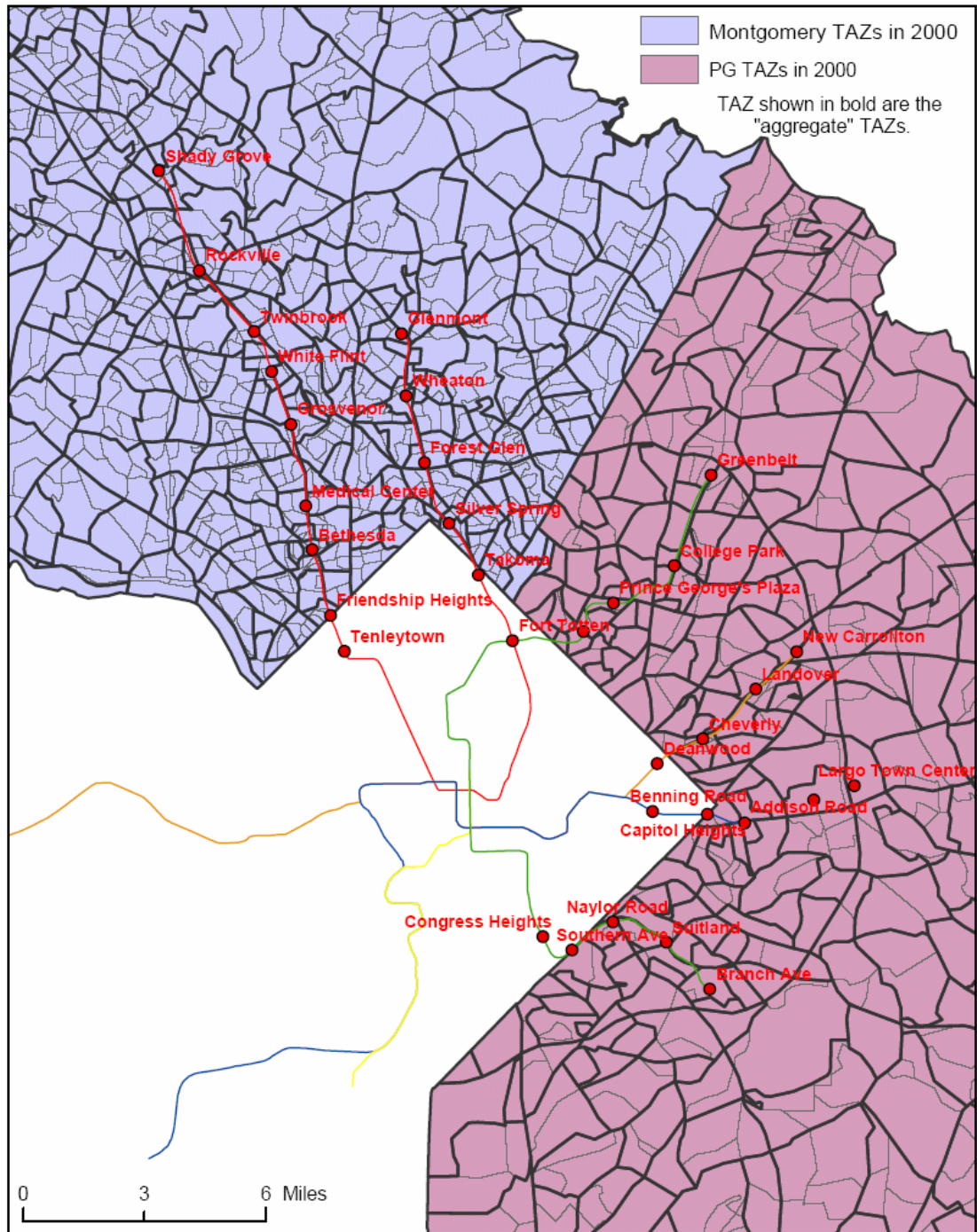


Table 3.2: Boardings by station and development potential¹

	Boardings				Growth earliest to 2000 ⁴	Year opened	Development potential in 1983	Parking spaces 2005
	1985	1990	1995	2000				
Montgomery Station								
Bethesda	5,011	7,572	7,501	8,469	0.69	1984	high	0
Forest Glen			1,898	1,860	-0.02	1990	high	596
Glenmont ²				4,619		1998	high	1,781
Grosvenor	2,618	3,794	3,438	3,551	0.36	1984	high	1,796
Medical Center	3,137	4,501	3,911	4,141	0.32	1984	government	0
Rockville	2,140	3,944	3,443	3,844	0.80	1984	low	524
Shady Grove ²	4,050	9,106	9,014	9,767	1.41	1984	low	5,467
Silver Spring	13,445	14,857	11,311	11,492	-0.15	1978	high	0
Takoma	5,195	6,227	5,204	5,595	0.08	1978	low	0
Twinbrook	2,354	4,515	3,778	3,863	0.64	1984	low	1,097
Wheaton			5,508	4,001	-0.27	1990	high	977
White Flint	2,199	4,333	3,695	4,050	0.84	1984	high	982
TOTAL	40,149	58,849	58,701	65,252				13,220
Prince George's Station³								
Addison ²	2,723	5,703	5,238	6,682	1.45	1980	low	1,268
Capitol Heights	2,317	2,922	2,099	2,324	0.00	1980	low	372
Cheverly	1,315	1,655	1,484	1,505	0.14	1978	low	530
College Park			1,504	2,709	0.80	1993	medium	530
Greenbelt ²			2,948	5,786	0.96	1993	high	3,399
Landover	2,940	3,856	3,192	3,477	0.18	1978	low	1,866
New Carrollton ²	5,695	8,786	7,670	8,742	0.54	1978	high	1,772
PG Plaza			2,391	3,389	0.42	1993	high	1,068
West Hyattsville			1,876	2,793	0.49	1993	high	453
TOTAL	14,990	22,922	28,402	37,407				11,258

¹ Ridership data provided by Robert Stedman in the Washington Metropolitan Area Transit Authority (WMATA). The data correspond to daily boardings obtained in May. The development potentials were assessed by MWCOG (1983). The number of long term parking spaces as well as the year opened come from the information submitted by the WMATA on their internet page www.wmata.com.

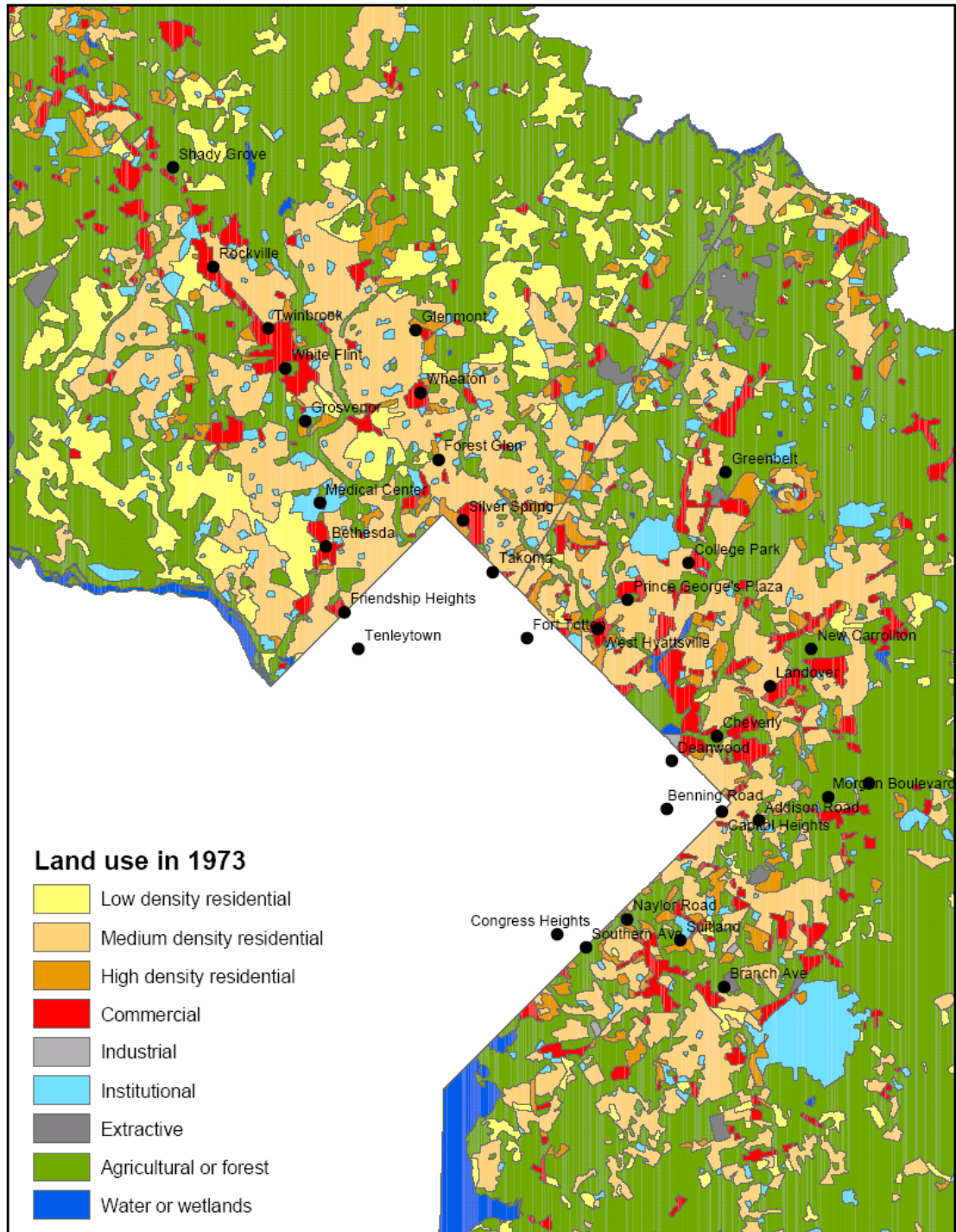
² Terminal station

³ In 2000 the following stations had not opened yet: Morgan Boulevard, Largo Town Center, Branch Avenue, Suitland, Naylor Road, and Southern Avenue

⁴ Some of the stations show decreases in the use, but mainly these can be explained by the fact that as new stations farther along the line came into operation some people using for example the Silver Spring station switched to using either Wheaton or Glenmont.

The stations were also built in different types of neighborhoods. Some were more residential while others had higher concentrations of land devoted to employment. Map 3.2 shows the land use in 1973 in the two counties.

Map 3.2: Land use in Montgomery and Prince George's Counties in 1973



Data from the Maryland Department of Planning, Maryland PropertyView CD

Differences can also be observed by examining the current patronage levels by station. It is evident that in Montgomery County the patronage of the network is higher. The average boarding per station in 2000 in Montgomery was 5,438 passengers per day, whereas for an average Prince George's station it was only 4,156 passengers. Part of these differences may be due to the fact that fewer people live or work near a station in Prince George's than in Montgomery. In 2000 there were 5,885 workers who resided in a traffic analysis zone (TAZ) that was within a mile of a station and 7,344 workers who worked in a TAZ that was within a mile of station in Prince George's County. The comparable numbers for Montgomery were 7,290 and 14,232, respectively. It is also true that the percentage of those living or working near a station and using the Metro for commute purposes was, in general, higher in Montgomery than in Prince George's (Tables 3.3a and 3.3b). In Montgomery County, over 20% of the workers residing in a TAZ within a mile of the Bethesda, Silver Spring, or Takoma stations and over 20% of the workers working near the Grosvenor, Takoma, and Glenmont stations commuted to work using the Metro. In Prince George's County no station was used by more than 20% of those commuting to or from areas within a mile of a station. Overall, in Montgomery 16% of those residing near a station and 8% of those working near stations favored the rail mode for commute. The comparable numbers for Prince George's County were 14% and 2%. However, even within a county there were differences: some stations in Montgomery County were used by a relatively small percentage of the potential workers for commute purposes.

Table 3.3a Share of Metrorail commuting by station

Montgomery Station	Within one mile ¹				Closest station ¹					
	Total population of a station	Total workers residing within a mile of a station	Pct of workers taking Metro whose residence is within a mile of...	Pct of workers taking Metro whose employment is within a mile of...	Total population	Total workers residing with the closest station being...	Total employment with the closest station being...	Pct of workers taking Metro whose residence is... is...	Pct of workers taking Metro whose closest station to employment is...	
Bethesda	16,335	9,845	31,575	0.20	0.06	35,150	19,150	37,585	0.13	0.07
Forest Glen	12,610	6,675	6,119	0.16	0.18	34,020	17,380	13,263	0.12	0.01
Glenmont ²	16,945	7,920	1,884	0.11	0.48	150,159	75,690	26,239	0.06	0.01
Grosvenor	14,105	8,249	5,710	0.17	0.25	33,470	18,244	27,129	0.11	0.01
Medical Center	9,425	5,285	26,580	0.18	0.03	31,470	15,635	30,909	0.10	0.04
Rockville	14,675	7,565	22,464	0.11	0.04	61,725	31,645	40,048	0.07	0.02
Shady Grove ²	4,545	2,430	9,730	0.17	0.04	299,387	164,691	135,431	0.05	0.01
Silver Spring	22,710	13,695	20,210	0.21	0.14	54,180	29,044	23,270	0.14	0.06
Takoma	8,055	4,490	3,235	0.22	0.30	21,185	11,035	6,570	0.15	0.02
Twinbrook	15,980	7,440	19,380	0.13	0.05	41,615	20,180	24,375	0.10	0.02
Wheaton	16,570	8,300	6,293	0.10	0.14	48,580	24,429	17,743	0.08	0.02
White Flint	9,915	5,580	17,605	0.14	0.04	26,515	13,865	23,300	0.09	0.02
TOTAL	161,870	87,474	170,785							

¹ Data calculated from 2000 CTPP files. In the first set of columns (within a mile) are included all people who work or reside in a TAZ that is within a mile of a station. In the second set of columns are included all people who work or reside in a traffic analysis zone (TAZ) whose closest station is the given station. Some of these people may be more than 20 miles from the actual station.

² Terminal station

Table 3.3b Share of Metrorail commuting by station

Prince George's Station ²	Within one mile ¹				Closest station ¹				
	Total population of a station	Total workers residing within a mile of a station	Total employment within a mile of a station	Pct of workers taking Metro whose residence is within a mile of...	Total population	Total workers residing with the closest station being...	Total employment with the closest station being...	Pct of workers taking Metro whose residence is...	Pct of workers taking Metro whose station to employment is...
Addison ³	12,610	5,245	855	0.19	199,770	97,684	74,516	0.06	0.01
Capitol Heights	6,190	2,335	1,605	0.16	20,730	8,530	7,450	0.12	0.01
Cheverly	10,335	4,700	6,135	0.15	15,080	6,990	11,520	0.13	0.01
College Park	9,660	5,395	21,045	0.11	34,955	16,190	27,980	0.07	0.02
Greenbelt ³	13,050	6,840	10,820	0.09	108,480	60,439	69,203	0.05	0.01
Landover	17,020	7,150	4,020	0.13	41,870	17,950	8,395	0.10	0.01
New Carrollton ³	13,190	6,195	12,480	0.14	152,349	79,565	53,169	0.07	0.01
PG Plaza	15,110	7,570	6,930	0.12	48,255	23,440	14,280	0.09	0.02
West Hyattsville	16,270	7,535	2,209	0.16	23,815	10,940	4,729	0.14	0.03
TOTAL	113,435	52,963	66,099						

¹ Data calculated from 2000 CTPP files. In the first set of columns (within a mile) are included all people who work or reside in a TAZ that is within a mile of a station. In the second set of columns are included all people who work or reside in a traffic analysis zone (TAZ) whose closest station is the given station. Some of these people may be more than 20 miles from the actual station.

² In 2000 the following stations had not opened yet: Morgan Boulevard, Largo Town Center, Branch Avenue, Suitland, Naylor Road, and Southern Avenue

³ Terminal station

Besides differences in station utilization, stations also had different perceived potentials for development (Metropolitan Washington Council of Governments, 1983). Based on this classification about half of the station areas were rated as highly developable in 1983 (Table 3.2). A greater number of stations in Montgomery County received this classification than in Prince George's County. However, the Prince George's stations that opened in the 2000's were not evaluated in the study. These differences imply that from the beginning, as assessed by the MWCOG, stations were expected to react differently to the opening of the subway. This variability is confirmed by Table 3.4 in which the population, employment and land use characteristics in the early 1970s are listed for each station.

Table 3.4: Characteristics of Metro station areas (in operation in 2000)

	Population density per hectare in 1970	Employment density per hectare in 1972	Pct of area in low density residential in 1973	Pct of area in medium density residential in 1973	Pct of area in high density residential in 1973	Pct of area in commercial activity in 1973	Pct of area in institutional uses in 1973
Montgomery Station							
Bethesda	30.95	150.39	0.29	0.38	0.03	0.20	0.04
Forest Glen	18.74	36.82	0.03	0.59	0.03	0.05	0.08
Glenmont ¹	27.56	8.82	0.01	0.65	0.10	0.04	0.07
Grosvenor	15.20	19.39	0.10	0.32	0.08	0.04	0.08
Medical Center	12.56	133.22	0.11	0.32	0.01	0.01	0.29
Rockville	19.13	51.73	0.17	0.42	0.03	0.26	0.08
Shady Grove ¹	1.12	17.84	0.11	0.00	0	0.14	0
Silver Spring	49.93	129.46	0	0.52	0.18	0.25	0.02
Takoma	37.88	54.38	0	0.78	0.07	0.05	0.07
Twinbrook	29.05	82.74	0	0.30	0.11	0.32	0.02
Wheaton	25.32	56.20	0.09	0.59	0.07	0.12	0.06
White Flint	8.08	37.75	0	0.31	0.02	0.39	0.06
<i>Average</i>	<i>23.10</i>	<i>65.65</i>	<i>0.07</i>	<i>0.45</i>	<i>0.06</i>	<i>0.15</i>	<i>0.07</i>
Prince George's Station							
Addison ¹	9.67	2.70	0.02	0.17	0.06	0.06	0.01
Capitol Heights	43.94	15.04	0.02	0.58	0.12	0.08	0.09
Cheverly	17.50	15.38	0.01	0.49	0.03	0.21	0
College Park	23.71	14.90	0	0.48	0	0.16	0.06
Greenbelt ¹	19.75	17.18	0.02	0.25	0.13	0.13	0.01
Landover	29.49	14.76	0	0.43	0.09	0.22	0
New Carrollton ¹	21.28	11.56	0.05	0.19	0.09	0.26	0.02
PG Plaza	32.39	24.03	0	0.61	0.03	0.18	0.05
West Hyattsville	45.31	10.63	0.01	0.38	0.19	0.16	0.09
<i>Average</i>	<i>27.00</i>	<i>14.02</i>	<i>0.01</i>	<i>0.40</i>	<i>0.08</i>	<i>0.16</i>	<i>0.04</i>

The station areas are determined as those TAZs that are within a mile of a Metro station. The data for population comes from the 1970 Census of Population; the employment data are from Green and James (1993); the land use data are from Maryland Department of Planning.

¹ Terminal station

Furthermore, there is variability in the characteristics of different neighborhoods surrounding the station areas. This variability is analogous to differences in the individuals receiving training. In order to reduce the variance within the observational units, I use traffic analysis zones (TAZ) as the unit of observation. In general, TAZs follow the U.S. Census geography and are aggregations of various census blocks, census block groups or at times of census tracts. State or local planners determine TAZ

boundaries and use them to determine commuting patterns in their jurisdictions (U.S. Census Bureau, 2001). In 2000 there were 1,272 TAZs in Montgomery County and 623 in Prince George's County. The size of a TAZ in Montgomery and Prince George's counties ranged from less than a hectare to 1,250 hectares.³⁴ The smaller TAZs were located in the more populous parts of the counties. The average TAZ is about 100 hectares. Given their size, they are small enough to be relatively homogenous in their characteristics. TAZs are also chosen as the observational unit given that it is the smallest geographic unit for which employment data in 2000 and 1990 are available from the Census Transportation Planning Package (CTPP). Map 3.1 shows the distribution of the 2000 TAZs for the relevant parts of the two counties. The boundary files for the TAZs come from the Bureau of Census' geographical databases.³⁵

Some TAZs are excluded from the analysis based on the amount of land in protected uses or in water/wetlands. The water bodies are based on the 1973 land use data from the Maryland Department of Planning and the protected land areas come from the Maryland Department of Natural Resources.³⁶ The protected uses included are: County parks, Department of Natural Resource's lands and federal lands. If more than 50 percent of land is in protective uses or if the TAZ is a body of water, the TAZ is not included in the analysis. This removes from the sample areas such as Andrew's Air Force Base. In all, thirteen TAZs in Montgomery County and twenty-one observations in Prince George's County are excluded by the above rule. Also, in 2000, a Montgomery

³⁴ A hectare is 10,000 square meters or 2.47 acres.

³⁵ The 2000 boundaries can be found at <http://www.census.gov/geo/www/cob/tz2000.html>, and the 1990 boundaries in the 1990 Census Transportation Planning Package databases.

³⁶ The 1973 land use data can be found on the Maryland PropertyView CDs, and the information on protected lands can be found at <http://dnrweb.dnr.state.md.us/gis/data/data.asp>.

TAZ with an area less than 0.05 hectares is excluded from the analysis because of its small size. The characteristics of the TAZs are discussed in Section 3.4 in more detail.

It must be stressed that in my analysis the impacts obtained are only average treatment impacts. That is, there will be variability in the impact depending on the individual characteristics of each station and neighborhood. These heterogeneous impacts cannot be further evaluated given the relatively small number of observations.

3.2.3 Temporal Impacts

As with any policy, results are unlikely to occur immediately. As emphasized in Chapter 1, it is necessary to allow sufficient time to pass between the application of the treatment and the impact analysis.

There are several reasons for expecting the impact on development patterns from the subway system to change through time such that the initial impact may be small but cumulative impacts in time considerably larger. First, it is possible that people do not realize the benefits (or disadvantages) when the system is new and need to experience them before the amenities (or disamenities) associated with the system are internalized. Second, with time, the system expands increasing the number of locations that are accessible via the system and increasing the accessibility benefits to existing stations. Third, if the system has had a positive impact on the locational decisions of households and employees, a self-enforcing cycle may occur. It may become increasingly more attractive to be close to a station since a greater number of potential trips can be made via the system independent of an increase in the number of stations. The increase in the

average daily boardings per station (Table 3.2) may be an indication of these effects. Based on the above, the impacts from a Metro system on development patterns may intensify through time. However, if there is a “ceiling” density (either due to preferences or inflexible zoning) then once that has been reached no additional development occurs, and only areas below this density would be expected to densify.

The above observations do not only apply to the system as a whole but also to individual stations. In any evaluation year different stations have been operating for varying lengths of time. Also the initial treatment dosages differ depending on when in the history of the system the station opened. A more recently opened station can be expected to have greater impact in the first x years than a similar station opened earlier, *ceteris paribus*. The more developed the subway network, the greater the potential benefits from any particular station.

In order to explore the impact of time on the development densities, I use information from 1990 and 2000 for the outcome measures for Montgomery and Prince George’s counties. This comparison across the two years gives an indication whether impacts indeed are growing. One potential problem with the comparisons is that between 1990 and 2000 additional stations were opened. Therefore, the two samples of stations are different. Even if the samples were restricted to the same set of stations, there would be variability in when the stations opened, causing the stations to be at different stages of the development process. To examine the impacts holding the opening year constant I

look at impacts on the western branch of the Red Line in Montgomery County.³⁷ All of these stations were opened in 1984 and thus changes in the degree of impacts are more closely correlated with the lapse of time than other factors. Such analysis is not possible for another set of stations, given the small number of stations that opened simultaneously in other years.

3.2.4 *Spatial Impacts*

The impacts of a subway station on employment and population in a TAZ potentially depend on the distance of the TAZ from the station. That is, areas (TAZs) closer to the station receive a higher dose of treatment than areas farther from the station. However, the impact on densities is not necessarily a linear one. It is possible that the degree of impact is highest at some intermediate distance (dosage).³⁸

Any treatment radius is, however, arbitrary. There are several possible consequences from improperly delineating the treatment area. If the radius for the treatment area is too small and does not cover all of the area affected, then the control group will be “contaminated” with treated observations. There will be spillover impacts to the control group TAZs. This contamination implies that any impact based on the analysis would be dampened. Similarly, if the radius for the treatment area is too large, the treated group is contaminated by controls and again the true impact would be

³⁷ This includes the stations areas surrounding Bethesda, Medical Center, Grosvenor, White Flint, Twinbrook, Rockville, and Shady Grove stations.

³⁸ For example, Bowes and Ihlanfeldt (2001) find that the greatest positive impact on residential housing prices from a rail station occurs between 0.25 miles and 0.5 miles from the station and not at the closest proximity to the station.

underestimated. The implication is that any impact observed is a lower bound of the true impact if the impacts are positive.

To explore the influence of distance, two different treatment distances are considered. In the first analysis, the treatment group includes all the TAZs that have their centroid within one mile of a station and the possible controls are those farther than one mile from a station.³⁹ One mile is chosen as the cut-off distance given that it is a reasonable distance to walk. A person in average physical condition walks a mile in about 15 to 17 minutes. This may be considered a reasonable time to reach the transit station (or possibly an upper limit).⁴⁰ This is also a distance that encompasses most walk-based trips to transit stops. O’Sullivan and Morrall (1996) find that the average distance walked to a suburban station in Calgary is 1.1 kilometers (0.68 miles). Marchwinski (1998), in a study in Pennsylvania, finds that 70 percent of those who walk or bike to a commuter rail station live within half a mile of the station, while 91% live within a mile of the station. The Metropolitan Washington Council of Governments, the planning agency for the Washington metropolitan area, uses one mile as the limit for a “long” walk in their traffic/mode choice simulations (MWCOG, 2004).⁴¹ Beyond that, nearly everyone will drive (or take other public transportation that may or may not connect with the Metro system) and thus the potential accessibility improvement by the Metro

³⁹ This definition implies that parts of the TAZs included in the treatment sample are beyond a mile of a station and there are some areas that are within a mile radius, but belong to a TAZ with its centroid farther than a mile from a station, that are considered in the control group. In general, the part of the TAZ that is misclassified is small. Any impacts derived are going to be a lower bound estimate of a true impact.

⁴⁰ Based on evaluations done with the Rockport Fitness Walking Test (for example, for college students see Byars, Greenwood, Greenwood, Simpson, 2003; and for healthy adults 40 to 79 years old see Pober, Freedson, Cline, McInnis and Rippe, 2002)

⁴¹ Obviously the actual possibility of walking (or biking) will depend on the sidewalk and road conditions (Saelens, Sallis, Frank, 2003), but these factors are not considered in the analyses. The implicit assumption is that pedestrian facilities are homogeneous within the treatment areas.

decreases. From Tables 3.1 and 3.2 it is clear that some subway riders drive to stations, especially to a terminal station with extensive parking facilities. However, those that drive, in theory, may locate anywhere in the region to take advantage of the network. If the cost savings are sufficient, some may choose to live farther out with the subway than had the subway not existed. They will locate farther away from the center, where land rents are lower, and will commute to a station with parking facilities. If this is the trend, no densifying impacts are expected.

In the second analysis only TAZs within a half mile radius of a station are considered in the treatment group. If the accessibility benefits vary with distance from the Metro station, so that locating closer to the station yields higher transportation cost savings, we would expect greater impacts in areas closer to stations than in areas farther away. Also, as discussed above, most people who do take transit services walk much less than a mile—supporting the use of a smaller distance. Since treatment impacts are believed to occur within a radius of one mile, I use as the possible controls those TAZs farther than one mile from a station even for the half-mile treatment sample: In these analyses the TAZs that are between a half mile and a mile are not included in either the treatments or in the controls.⁴² This designation also reduces the impact of any spillovers, since no controls can be directly adjacent to the treated observations considered in the sample. However, it does not eliminate possible spillovers if the station area affects development beyond the one-mile radius. To test for possible spillover

⁴² Considering TAZs between half a mile and one mile as possible controls would change the interpretation of the results. In this case, the implicit assumption would be that the treatment area extends to only those TAZs within half a mile of a station. Since a fair number of people beyond the half a mile radius walk to the station this assumption seems restrictive and spillover impacts are possibly substantial.

effects in the one mile radius sample, I restrict the potential control TAZs to those that are farther than 1.5 miles from a station.

To further analyze the impacts at differing distances, I apply the multiple treatment framework with three different treatment groups. The first group is comprised of those TAZs within half mile from a station. These are the areas that received the largest dose of the treatment. The second group is composed of those stations that are within half a mile and a mile from a station. The third group includes the remaining TAZs, those that are farther than a mile from a station.

3.3 Outcome measures

As discussed in Chapter 1, subway stations may affect the location decisions of households or firms if the transit network reduces the costs of transport. If the benefits are sufficiently large, the transit network may alter the distribution of population and employment in the region, leading to higher densities in areas where the benefits from the transit network are large. In order to capture these potential changes in the location decisions of households and firms, I develop three different sets of measures. First, to measure the impact on the location decisions of firms, I use employment density. Second, to measure the impact on households, I use not only population density, but also measures that capture the socio-economic characteristics of the population. Specifically, I use average family income as well as the percentage of households belonging to a minority group. Third, given that some Metro areas may develop into commercial

centers and others into residential areas, I also include a measure of overall development density. Following is a discussion of the specific measures used in the analyses.

3.3.1 Employment

To capture commercial density development, I use employment density of the TAZ. Employment density indirectly measures the willingness of firms to pay for a particular location. As discussed in Chapter 1, when the resulting savings in transportation costs or the transaction costs with customers and suppliers are sufficiently large, firms will locate near transport nodes. In the aggregate, these decisions increase the price of land and development density. Employment density is calculated as the total number of employees in a TAZ per hectare. The land area used as the denominator is the effective net land area, which is the land area of the TAZ adjusted for water and wetlands as well as for the area in protective uses. Included in protective use areas are national and state parks, such as the C&O Canal National Historical Park. The reasoning behind excluding these areas from the density calculations is that, in theory, they will not be converted into developed uses and are thus outside of the development process modeled here. The employment figures come from the Census Transportation Planning Packages (CTPP) based on the 1990 and 2000 Censuses. The CTPP gives population and employment characteristics at the traffic analysis zone level. The geographical files for protected lands come from Maryland Department of Natural Resources.

3.3.2 Population and socio-economic characteristics

Several measures are used to capture the household's location decision. As with to employment density, population density measures indirectly the willingness of

households to pay to locate in a particular area. It also directly measures the distribution of population within a metropolitan area. The measure is based on the 1990 and 2000 CTPP. Again the land area is adjusted for lands in protective uses and federal facilities such as Andrew's Air Force Base and the National Institutes of Health.

As an alternative set of measures of population density I use the dwelling unit density and changes in this measure. These measure the distribution of the housing stock. The first dwelling unit measure I use is the dwelling density in the outcome year. The second measure used is the change in dwelling units from 1970 to 1990 or from 1970 to 2000, and the third is the change in dwelling units in the decade prior to the evaluation, or from 1980 to 1990 or from 1990 to 2000, for outcome years 1990 and 2000, respectively. The last two measures capture the construction rate of new housing units. These measures are constructed from the Maryland PropertyView database and allow me to look at changes within a particular year's TAZ boundaries. The Maryland PropertyView database has information based on the tax assessment files for all of the properties in the State in a geo-referenced format.⁴³ I also calculate the dwelling density measure using information from the Census Transportation Planning Packages. The TAZ boundaries themselves change from one Census year to another. Thus, it is not possible to construct the change variables from Census data.

⁴³ Whereas the dwelling density measure for 2000 based on the PropertyView information is very accurate, the number of dwelling units in the prior periods are most likely undercounted. PropertyView has only the most recent information for any lot, any dwelling units that existed prior to the "current" one are lost from the database. The changes that are not captured in the system include a lot with a single-family house being converted into apartment buildings or other commercial uses. Given that the missing dwelling units are single-family houses most likely converted to higher density use, the underestimation in pre-2000 years is most likely not sizable and will not significantly affect the results.

Third, in order to evaluate whether any impacts are observed on the socio-economic makeup of the station areas, I construct two different measures—average family income and the share of population belonging to a minority group. The first is the average family income in the TAZ in either 1989 or 1999 from the Census Transportation Planning Packages. The second is the percentage of population in the TAZ who consider themselves members of a minority (non-white) group. This information is also obtained from the Census Transportation Planning Packages. With these measures I can address whether or not income determines the choice set of housing options near a station, which in turn, throws light on whether prices are likely to have increased such that lower income people cannot afford to live in the station areas.

3.3.3 *Overall Density*

Given that land dedicated to employment uses will not be available for residential uses and vice versa, it is possible that we do not observe differences in employment or population when these measures are considered separately. For example, it is possible that some TAZs have no population but are intensively built up as commercial centers. In this case, in the analysis of population, the TAZ would enter as a “zero” density TAZ and in the employment analysis it would enter as a densely built up TAZ. In order to overcome the problem of “specialization” of TAZs, I use a measure to capture the overall development density of the TAZ. The measure is the sum of the number of employees plus the number of dwelling units in the TAZ divided by the land area. That is, the overall development density is given by $(e + dw)/area$ where e is the number of

employees in the TAZ, dw is the number of dwelling units in the TAZ and $area$ is the land area of the TAZ.⁴⁴

3.4 Explanatory variables in the propensity score estimation

Variables used to capture the station siting decision process described in Section 3.1 include information on the location of the TAZ within the metropolitan area (distance to the White House), on pre-Metro population (population density as well as the mean family income, race, and the percentage of dwelling units in apartments), on pre-Metro employment (employment density), on historic land use (percent of area in TAZ in agricultural uses or in forests, percent of area zoned for low density residential uses and percent of area zoned for high density residential development), on recent growth (percent of 1970 dwelling units built after 1960), and on accessibility (distance to closest intersection of major roads). Table 3.5 gives the summary statistics for the variables constructed. Following is a more detailed discussion of the variables used to construct the propensity score.

⁴⁴ The correlation coefficient of the overall development density measure and employment density is 0.93 and 0.90 for Montgomery County and Prince George's County, respectively. The correlation coefficient between overall development density and dwelling unit density is 0.63 and 0.20 for Montgomery County and Prince George's County, respectively. The correlation between employment and dwelling unit densities is 0.32 and -0.01 for Montgomery and Prince George's counties, respectively.

Table 3.5: Descriptive statistics of initial conditions

	Montgomery County		Prince George's County	
	Mean	Std. Dev	Mean	Std. Dev.
Pct of population white in 1970 ¹	0.92	0.08	0.84	0.18
Mean household income in 1969, (\$1000) ¹	14.23	8.27	13.38	2.28
Mean income (\$1000) squared in 1969	271.04	240.64	163.19	85.34
Population density in 1970 (people/ha) ²	11.95	22.48	12.16	18.26
Distance to closest intersection (km) ³	1.59	1.12	1.76	1.23
Distance to the White House (km) ⁴	25.54	11.16	19.14	6.25
Pct of area in agricultural use or forest (adjusted for protected lands) in 1973 ⁵	0.44	0.41	0.52	0.35
Pct of dwelling units built post 1960 in 1970 ⁶	0.50	0.40	0.48	0.37
Pct of dwelling units in apartments 1970 ⁶	0.06	0.21	0.13	0.27
Pct zoned residential high density 1961 ⁷	0.02	0.10	0.05	0.12
Pct zoned residential low density 1961 ⁷	0.92	0.21	0.76	0.35
Employment density in 1972 (workers/ha) ⁸	5.74	15.81	3.57	7.67
Number of TAZs (with complete information)		1265		591

¹ Based on tract level averages from the 1970 Census of Population

² Based on tract level averages from the 1970 Census of Population re-distributed using dwelling units information from the Maryland PropertyView database

³ Distance from the centroid of the tract to the closest major intersection as determined by 1971 Maryland highway maps

⁴ Distance from the centroid of the tract to the White House

⁵ Based on a Maryland Department of Planning land use map for 1973

⁶ Based on information extracted from the Maryland PropertyView 1999 and 2002 databases

⁷ Based on a MNCPPC (1961) zoning map

⁸ Based on Green and James (1993) employment data, redistributed to the TAZs using land use data

3.4.1 Location of the TAZ

The location of the TAZ in the metropolitan area is calculated as the distance of the centroid of the TAZ to the White House. The distance is the Euclidian distance

joining the two points and it is calculated using ArcGIS 8 software. These distances are calculated for both the 2000 and 1990 TAZ centroids.

3.4.2 Population and socio-economic characteristics

The first measure to capture the characteristics of the people in the pre-Metro period is the population density of the TAZ in 1970. The population density measure is calculated using 1970 Census of Population tract level data.⁴⁵ I use information from the Maryland PropertyView databases to re-distribute the tract level data to the TAZs.⁴⁶ By overlaying the TAZ (either 2000 or 1990) and tract boundaries on the PropertyView data each property can be assigned to its respective TAZ and tract. Then the number of dwelling units per TAZ-tract is calculated using the year built information as well as the number of units for apartments for those structures that are multi-family dwellings.⁴⁷ Depending on the tax exemption code, some of the properties that appear in the database were excluded from the dwelling unit calculations. For example, properties used for health care facilities or educational institutions are not included given that they do not represent permanent residences. Appendix 3.1 lists in detail the categories excluded.

In Maryland PropertyView the year built information is less reliably filled in for commercial structures including apartment complexes. In order to obtain a more complete dataset, the information in the database is supplemented by information provided by the Montgomery County Department of Park and Planning and by Prince

⁴⁵ These data come from the Neighborhood Change Database by GeoLytics, Inc.

⁴⁶ The Maryland PropertyView databases used for Montgomery County are those of 1999 and 2003. For Prince George's County I use the databases for 1999 and 2004.

⁴⁷ Each condominium unit appears as a separate account but with the same geographical location as other units in the same structure.

George's County Planning Department.⁴⁸ The year built information provided is matched with the PropertyView database using the tax assessment account numbers for Montgomery County and a combination of name and address information for Prince George's County. Also it was necessary to check for multiple entries for the same property in the Prince George's files. At times a different tax assessment account number is given to the same property and thus it may appear more than once in the database. To avoid counting the apartment units more than once, it was necessary to correct the repeat observations based on their address and legal name information.

After knowing the tract-TAZ location of each residential property built before 1970, the total population of the tract is distributed to the TAZs based on the share of the tract's dwelling units in a particular TAZ. When a TAZ falls in two different tracts the TAZ population is obtained by summing over the population assigned to the TAZ from the different tracts. To obtain the density measure, the total TAZ population is divided by the net land area of the TAZ.

The three measures included to capture the pre-Metro socio-economic characteristics of a potential site, and therefore the expected "profitability" of a Metro station placed at this site, are the percentage of the population that is white and the mean household income, both based on the 1970 Census, and the dwelling units in apartments, based on the Maryland PropertyView database. For the first two measures the information is only available at the tract level and since it is difficult to re-distribute the

⁴⁸ The Montgomery County data were kindly provided by Wayne Koemple and those of Prince George's County by Philip Taylor.

Census tract level information (as done for the population density measure) each TAZ receives the tract's average value. If a TAZ is located in more than one tract, then the average value is calculated using as weights the percentage of the TAZ in each tract.⁴⁹ For the average income per family an additional correction is made. If there are no housing units in the TAZ the average income per household assigned is zero to reflect the fact that there was no current market for commuters in the particular TAZ.

To add more disaggregated information on the income of households, I include a third measure, the percentage of dwelling units in apartments. The measure crudely captures the ratio of owner versus rental units. A higher share of apartment units may make an area more attractive given that there is a potentially larger pool of riders. However, if the income of those living in the apartment units is low and residents prefer using the bus, they may be considered a non-profitable segment of the population and not be served by the Metro.

3.4.3 *Employment*

To capture the pre-Metro employment situation I estimate the employment density for each TAZ. The employment data are those used by Green and James (1993) in their analysis on the development impacts of the rail stations in the Washington metropolitan area. Their data are for aggregated TAZs and these are re-distributed to the disaggregated TAZs using the 1973 land-use information from the Maryland Department

⁴⁹ So for example, if a TAZ has 30% of its land area in tract A and 70% in tract B, the TAZ's calculated percentage of the population white would be $0.3*wA+0.7*wB$, where wA is the percentage of the population white in tract A and wB is the percentage of the population white in tract B. The mean household income is calculated the same way.

of Planning.⁵⁰ To obtain the employment density measure, I first calculate the share of each disaggregated TAZ in land-uses that are predominantly employment related (commercial, industrial, and institutional). The employment for the TAZ is then distributed using the share of the TAZ in the predominantly employment related activities as weights. Given that there is also some employment in residential areas, in the final employment density calculation five percent of the total employment is distributed to residential land use areas and 95% to employment related land areas. The above estimated employment is then divided by an adjusted net usable land area to obtain the employment density in 1972. The adjusted net usable land area includes some of the federal lands that are predominantly employment oriented. These areas include the areas belonging to the National Agricultural Research Center, the various military facilities of the region, the National Institutes of Health, and the Walter Reed Hospital.⁵¹

3.4.4 *Land use and zoning*

To capture the land use and zoning in the Maryland counties, I use a digitized Maryland Department of Planning map for 1973 and a digitized 1961 paper map of the Metropolitan-National Capital of Planning and Parks Commission for zoning. The 1973 land uses were grouped into nine general uses: residential low-density, residential medium-density, residential high-density, commercial, industrial, institutional, agricultural and forest, extractive, and water or wetland. Appendix 3.2 shows the make-

⁵⁰ That is, the TAZ nomenclature follows the general pattern of 3 numbers and 2 letters. It is possible to generate “aggregate” TAZs by joining the areas with the same 4 characters into one larger area.

⁵¹ To illustrate the procedure assume that the aggregate TAZ A is further divided into two areas TAZ A1 and TAZ A2. Also assume that A1 has 70 hectares in employment related uses and A2 has 30 hectares in employment related uses. Furthermore, both A1 and A2 have 50 hectares in residential uses. If the total employment in the TAZ A is 100 workers, using the algorithm to re-distribute employment into the disaggregate TAZs A1 has an employment of 69 workers and A2 has 31 workers yielding employment densities of 0.58 and 0.39 workers per hectare, respectively. In Map 3.1 the aggregate TAZs are shown with darker borders.

up of each aggregate group. I calculate, based on the TAZ boundaries, the percentage of the total TAZ land area in each of the uses. In the propensity score analyses I use as a proxy for the amount of completely undeveloped land the percentage of the TAZ land area in agriculture and forest uses. Again, to get a better idea of land potentially available for development I subtract from the measure the lands in protective uses.

The 1961 zoning categories are group into six general zones: residential low-density, residential medium-density, residential high-density, commercial, industrial, and rural. For the analyses the residential low-density and rural zoning categories are further aggregated into one single zoning category.⁵² Also I use the percentage of area zoned for high-density development. Given that each County designed its zoning categories with slightly different definitions, the two sets of zoning categories are not exactly the same.

3.4.5 Past growth

To capture past growth trends, I use the age structure of the housing units. I calculate the percentage of existing dwelling units in 1970 built in the 1960s. This information is calculated from the Maryland PropertyView data.

3.4.6 Accessibility

The accessibility measure used for the study is the distance from TAZ centroid to the closest major intersection. In order to identify major roads, Maryland State road maps from 1971 are used to determine the principal roads. These roads are for the most part comprised of the Federal and State Highways. A current digitized Maryland State

⁵² For Prince George's County the 1961 zoning map does not cover the whole County and I lose thirty-three TAZs that are located in the southern part of the County. Given that this area was mainly in agricultural and forest uses in 1973, these TAZs are unlikely to provide much information for the counterfactual.

Highway Administration map is then adjusted to reflect the principal roads in 1970.

Based on this map the intersections between major roads are identified and the distance from a TAZ to the nearest intersection is calculated.

Chapter 4

Case Study: Montgomery County

This chapter applies a kernel matching estimator to determine the impact of the Metro stations on development densities in Montgomery County. First, I describe the characteristics of Montgomery County prior to the opening of the Metro stations to gain a better understanding of the conditions in which the network was opened. Second, I analyze the estimated propensity scores based on the TAZs boundaries for both 1990 and 2000. Third, I report the treatment impacts on employment, population and overall development obtained via Epanechnikov kernel matching, focusing on the spatial and temporal nature of these impacts.

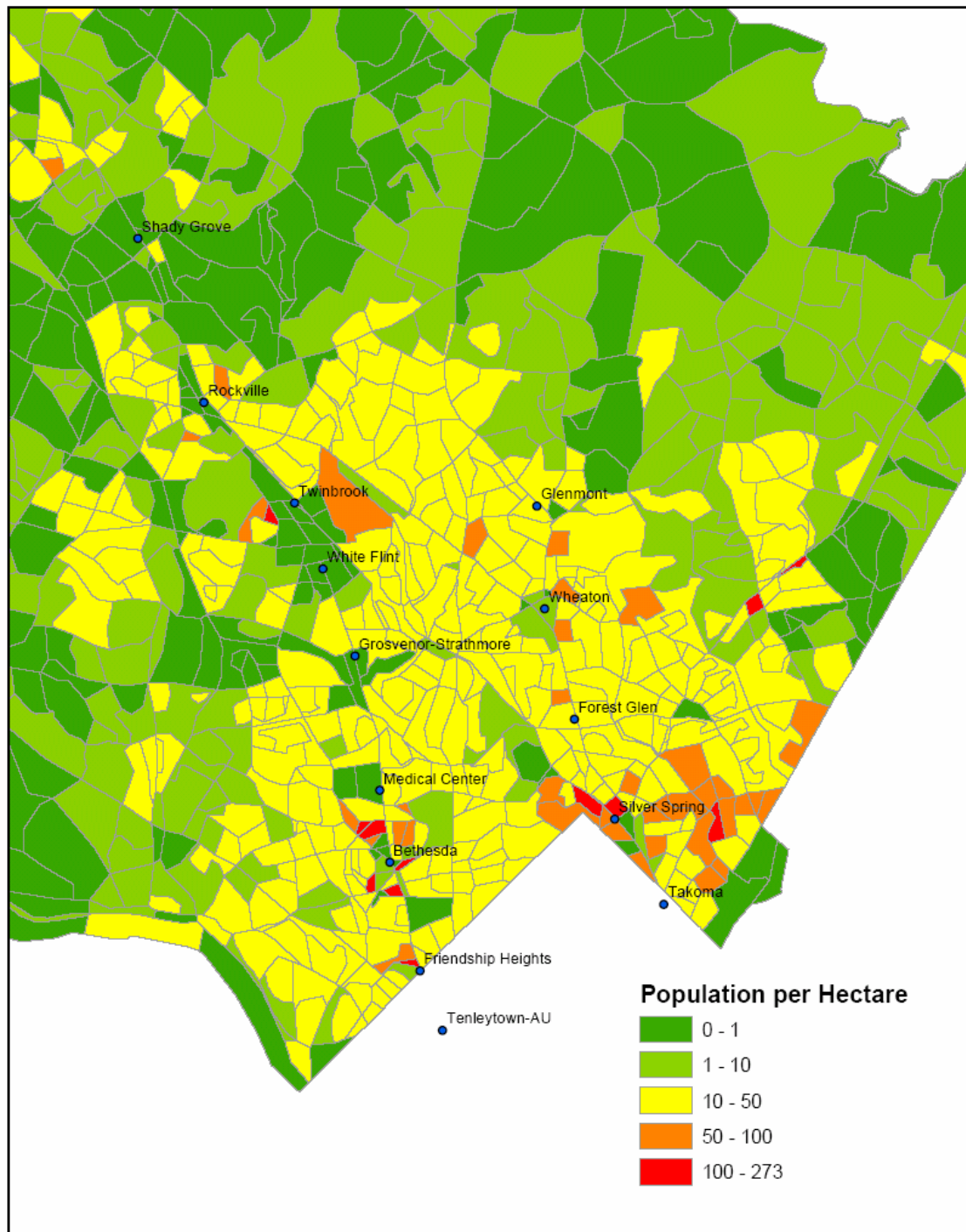
4.1 Initial Conditions in Montgomery County

Before analyzing the impacts of the Metrorail in Montgomery County it is useful to understand the differences that existed in the overall land use conditions prior to the system. The analyses underline the fact that the station areas were quite different in the early 1970s. Also, it will be evident that the TAZs near Metro stations were quite different from an average TAZ in the County, emphasizing the endogeneity in the station location decision.

In terms of the characteristics of the County in the early 1970s, most of the trends are as expected. Population density is higher closer to the District (Map 4.1). There are, however, pockets of low population density areas relatively close to the District line and

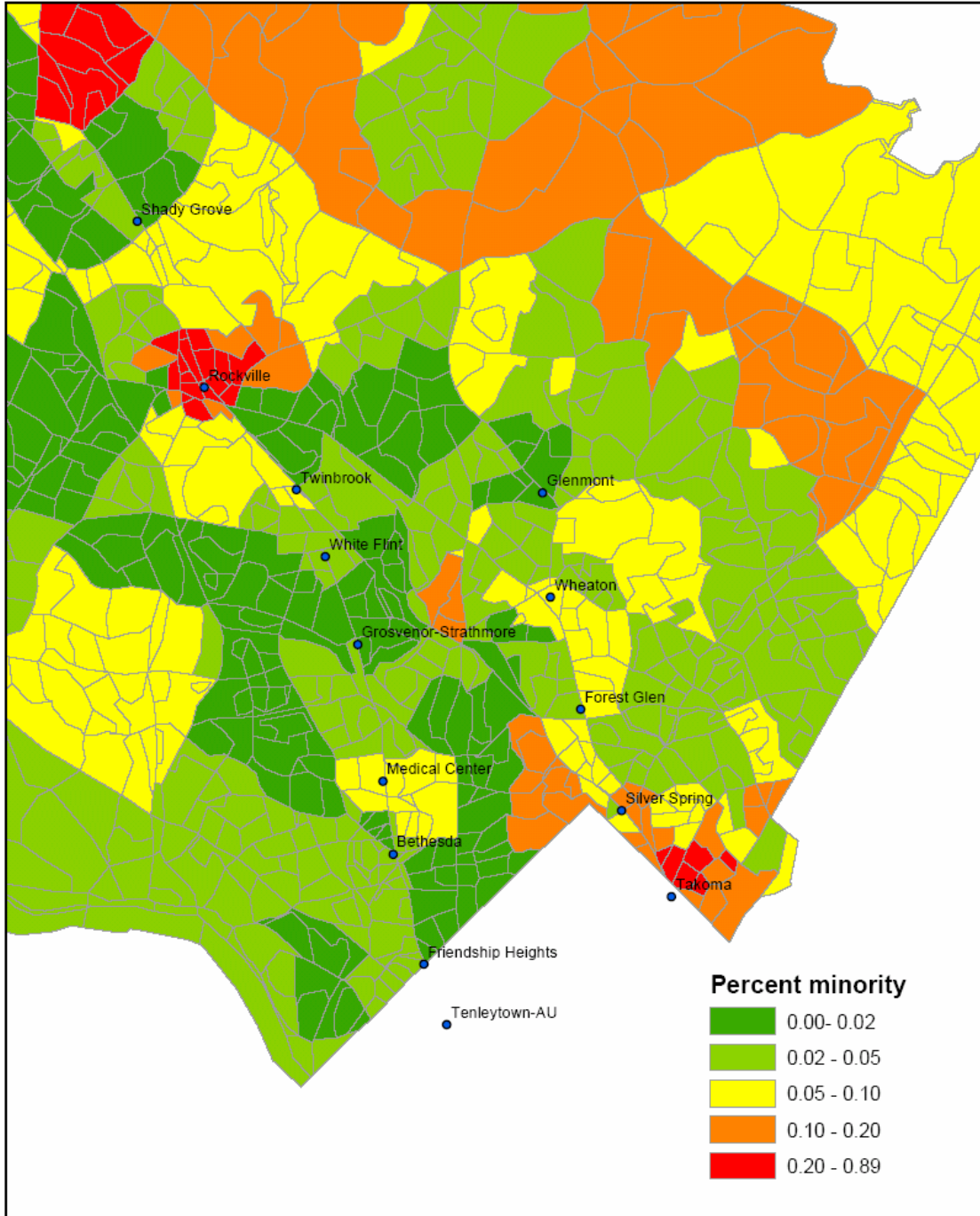
more densely populated areas further away. Many of the very highly populated TAZs are close to a Metro station but high population density was not interpreted by the MWCOG (1983) as an impediment to future growth potential (Table 3.2). For example, both Bethesda and Silver Spring were highly populated and also were rated as having high future development potential. It is also clear that there are TAZs close to Metro stations—such as Shady Grove, Twinbrook and White Flint—that were relatively sparsely populated. Of these the first two were categorized as having low development potential and White Flint as having high development potential. In terms of identifying possible controls, the TAZs in the northwestern part of the County had population densities of less than 10 inhabitants per hectare and should provide relatively weak controls for the typical treatment TAZ.

Map 4.1 Population Density in Montgomery County, 1970



Based on 1970 Census of Population; tract level data are redistributed to TAZs based on Maryland PropertyView

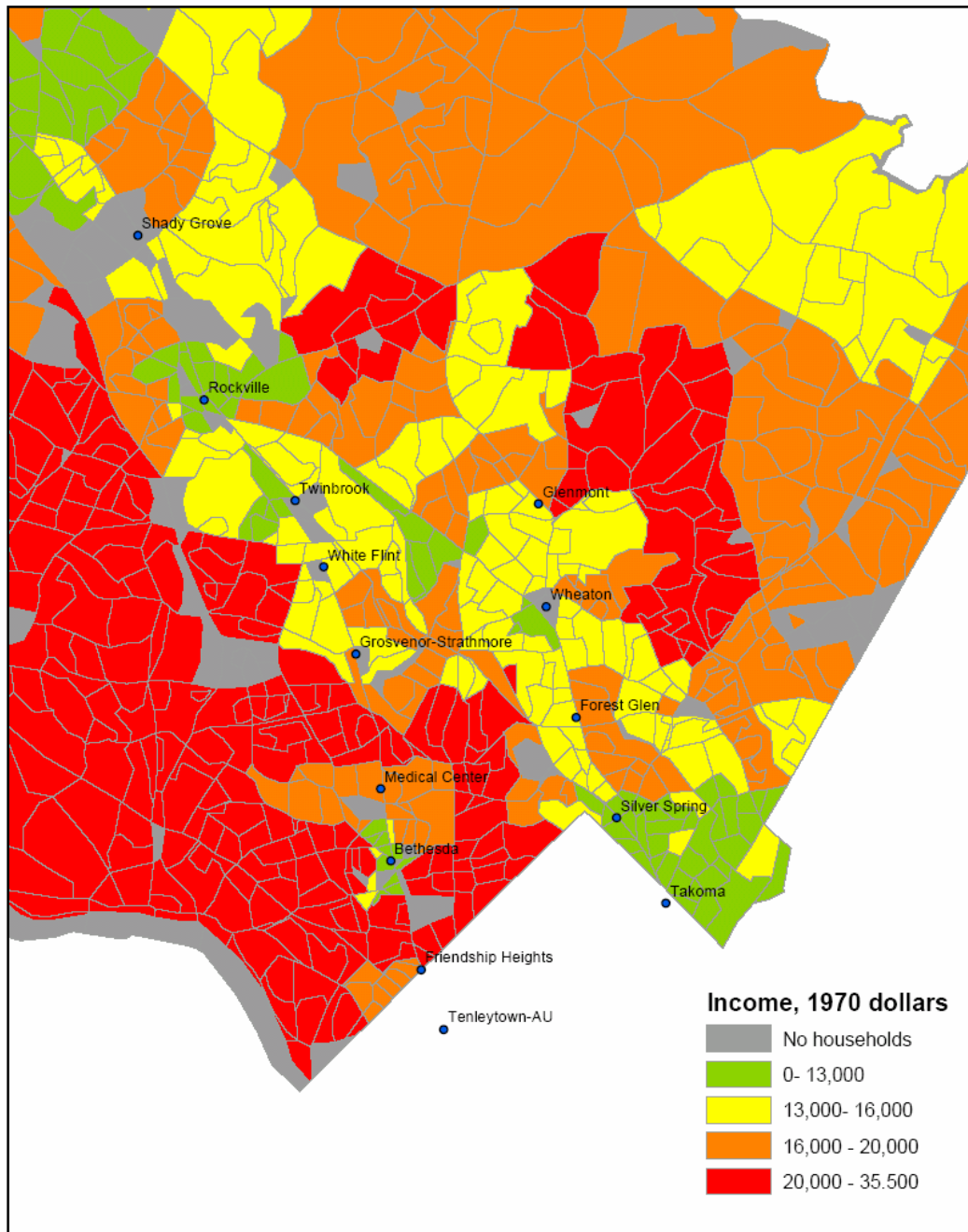
Map 4.2 Percentage of Population Belonging to a Minority in Montgomery County, 1970



Based on the 1970 Census of population

Map 4.3

Mean Household Income in Montgomery County, 1970



Based on the 1970 Census of population tract level data

In 1970, Montgomery County was predominantly white. On average less than 10% of the population in a TAZ belonged to a minority group. Map 4.2 shows the

percentage of minorities in each TAZ based on 1970 tract level Census data. There were two station areas that had a more substantial minority population: Rockville and Takoma. The other tracts with a higher concentration of minorities were in the rural northwestern part of the County. Map 4.3 shows the income distribution in Montgomery in 1970. The more rural areas of the County had lower average household incomes, and there were more uninhabited TAZs with zero incomes.⁵³ The highest average household incomes were located in the southwestern part of the County, largely beyond a mile from the closest Metro stop. As expected, the planners did not target the highest income households, but even so the stations were placed in relatively wealthy neighborhoods. These are areas in which the population had the means to use the service. It is important to acknowledge that both of these socio-economic measures are based on tract level averages with many TAZs grouped together in a tract. In reality there probably were differences among the TAZs within a tract, but unfortunately this diversity cannot be captured.

As with population density, many of the high-density employment TAZs were within a mile of a future Metro station (Map 4.4). Yet again, many of the stations areas had very low employment densities and within the controls there were several TAZs with moderately high employment densities. In many cases, the TAZs with high population densities had low employment densities and those with high employment densities had low population densities, suggesting the separation of land uses. There were, however,

⁵³ In the case of mean household income, some TAZs have “no income” given that they had no dwelling units in 1970 and thus the estimated TAZ population is zero. As the mean income proxies for the purchasing power in the TAZ, these TAZs are included in the analyses with zero average incomes. Some of zero income TAZs were exclusively non-residential developed uses, others were undeveloped. These TAZs occur both in the treatment and well as the control areas.

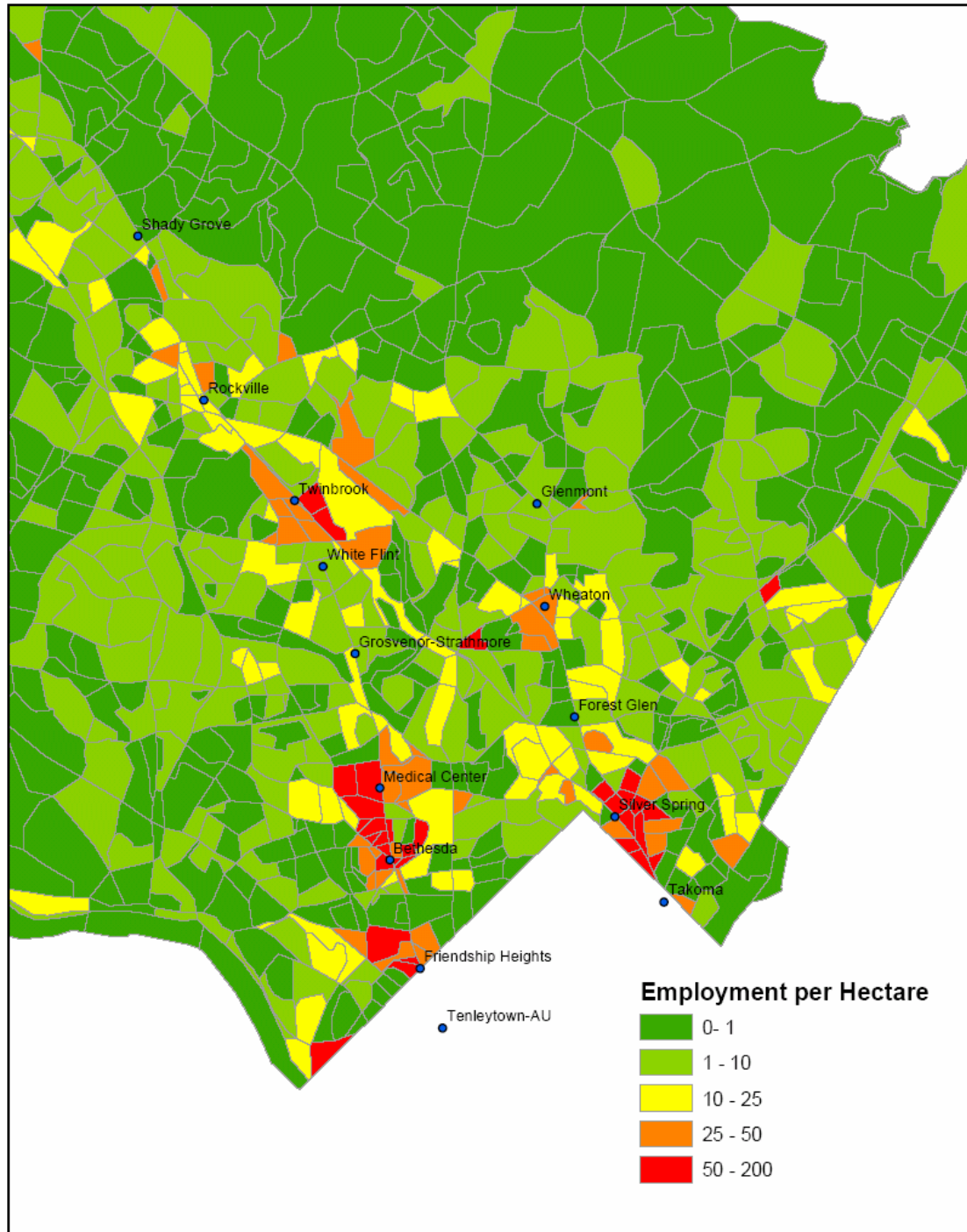
areas such as Shady Grove, White Flint and Grosvenor with both low employment and population densities and TAZs near Bethesda and Silver Spring that had both high employment and population densities.

The zoned land use in 1961 (Map 4.5) and actual land use in 1973 (Map 4.6) were both fairly non-uniform with varied zoned and actual land uses in the more urban portion of the County. The northern part of the County was mainly agricultural with small centers of commercial activity. In 1961 much of the land area within 19 miles of the White House was zoned for low residential and by 1973 the land use was quite mixed although medium density housing dominated the landscape.⁵⁴ There was high-density housing as well as commercial areas well beyond the one-mile radius from a Metro station. These two maps also suggest that zoning is endogenous. As development needs change, zoning also changes. That is, zoning in 1961 does not necessarily reflect the land use in 1973. Areas that were zoned rural or low-density housing may a decade later be in high-density residential, commercial or institutional uses.

The above distributions and descriptions of the initial conditions suggest that there are TAZs within the County that have similar characteristics to the treatment TAZs. It is also highly likely that for some treatment TAZs, especially those with very high employment densities, there are no control TAZs with similar characteristics. If so, this would suggest the need to use some subset of the treatment TAZs so as to homogenize the two groups for more robust results.

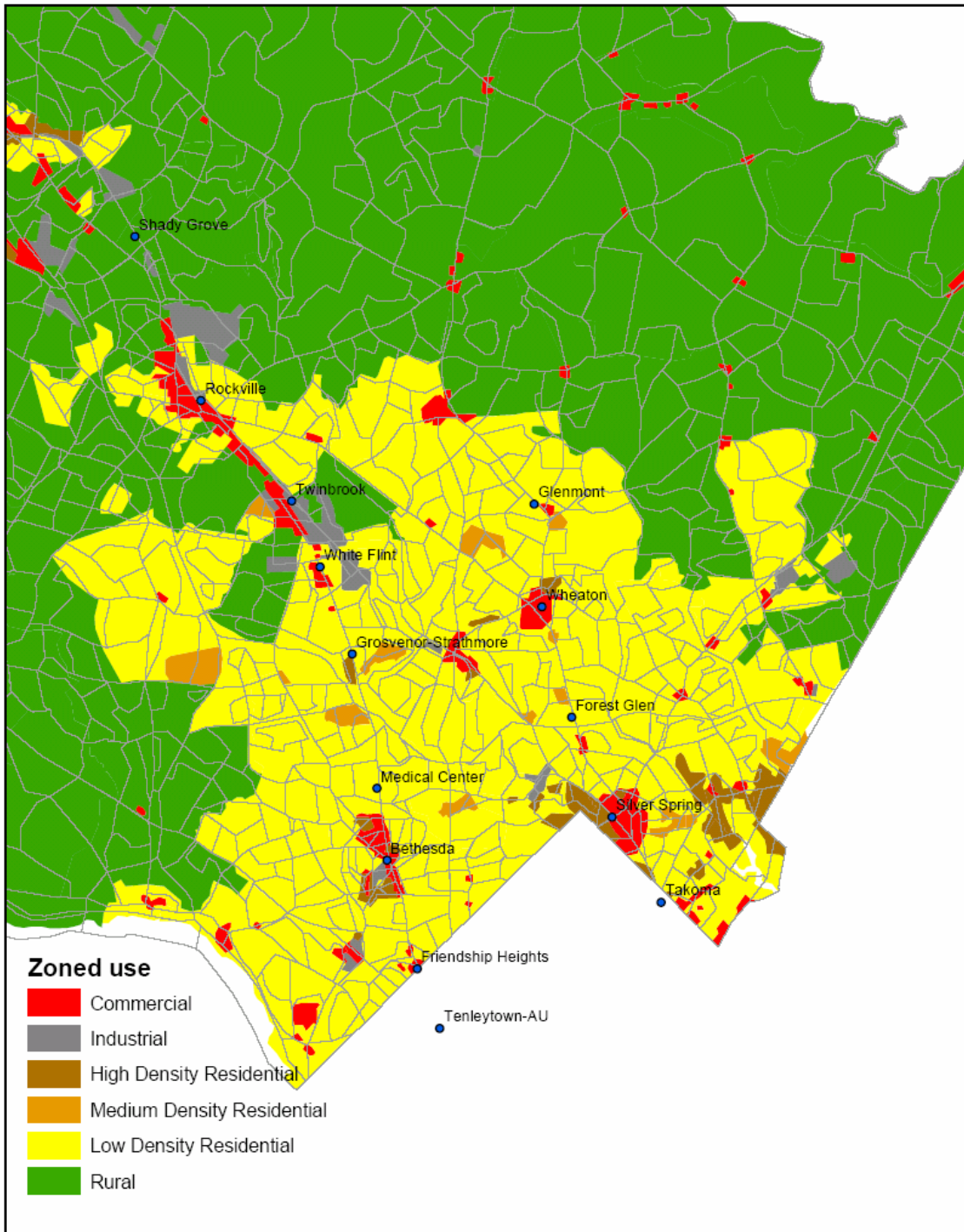
⁵⁴ The farthest station (Shady Grove) is about 17 miles from the White House.

Map 4.4 Employment Density in Montgomery County, 1972



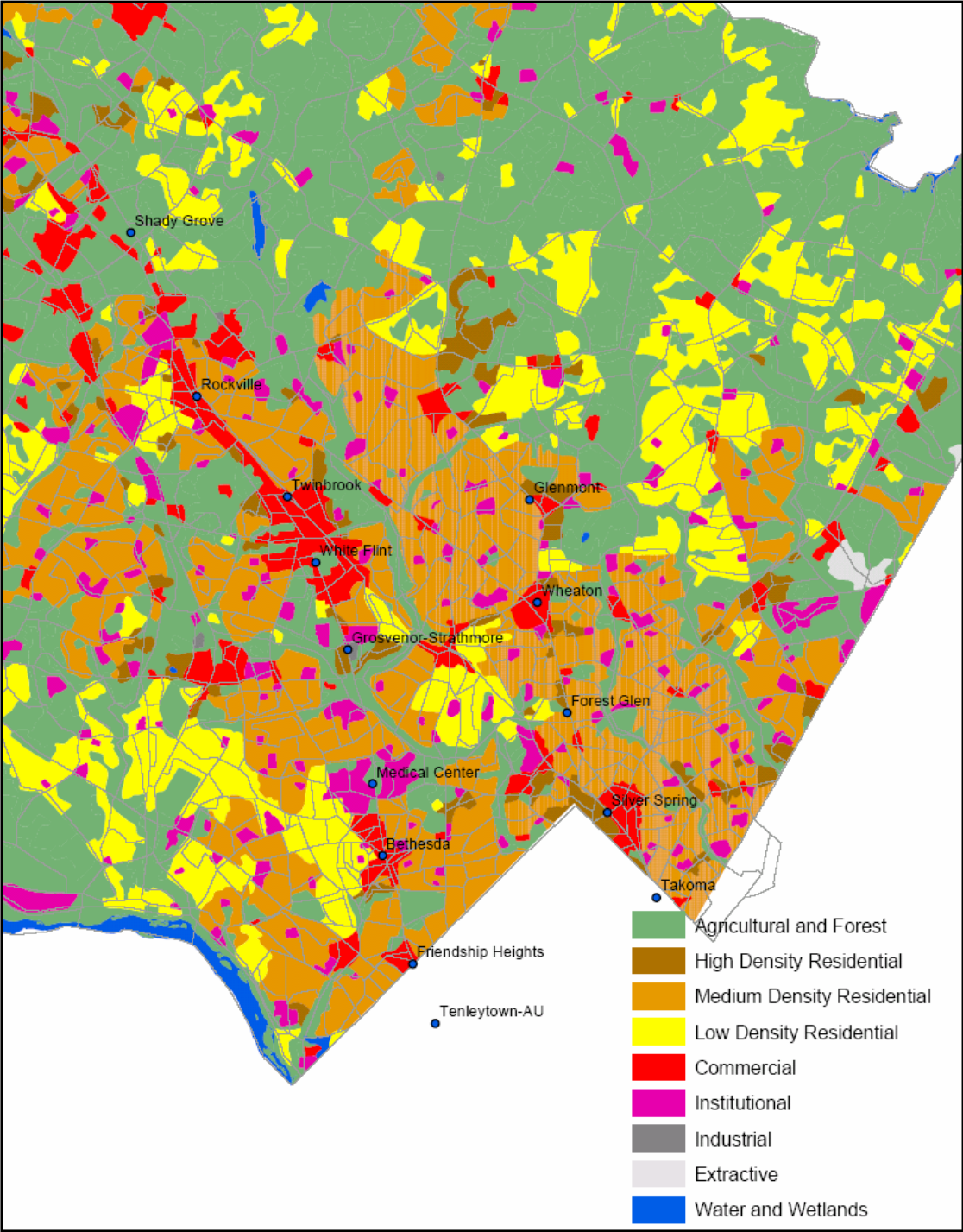
Based on 1972 aggregate TAZ level data from Green and James (1993)
Disaggregated TAZ level data are obtained by redistributing employment using land use information.

Map 4.5 Zoning in Montgomery County, 1961



Based on general zoning map in "Wedges and Corridors", MNCPPC, 1962

Map 4.6 Land Use in Montgomery County, 1973



Based on Maryland Department of Planning land use map, 1973

Looking qualitatively at the station areas by opening year highlights some of the differences in initial conditions depending on when the planned station opened. Table 4.1 presents the initial conditions for four different sets of station TAZs. The areas around the stations that were built in 1978 (the lower part of the eastern branch) are slightly different in their initial characteristics from the rest of the station areas. The stations built in 1978 were built in areas with a relatively higher 1970 population density, lower average household income, the TAZs were closer to the White House, with a higher percentage of apartment units in 1970 and a higher percentage of the area zoned for this type of development in 1961. Furthermore, the estimated initial employment density in 1972 was much higher. That is, the first stations that opened were in areas with more developed land use. The TAZs that are within a mile of a station that opened in 1984 (corresponding to the western branch) and in the 1990s (upper part of the eastern branch) are relatively similar. These station areas had on average more space for new denser development. The only difference is in the estimated 1972 employment density which is higher for the 1984 stations than the 1990s stations.

Table 4.1: Descriptive statistics of initial conditions around Montgomery metro stations by year opened for 2000 TAZs

	Station areas opened in 1978		Station areas opened in 1984		Station areas opened in the 1990s	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Pct of population white in 1970 ¹	0.88	0.07	0.94	0.08	0.95	0.02
Mean household income in 1969, (\$1000) ¹	12.35	2.71	13.76	7.63	14.01	5.06
Mean income (\$1000) squared in 1969	159.71	73.75	247.24	201.11	221.34	97.35
Population density in 1970 (people/ha) ²	45.07	41.93	24.22	39.02	25.70	18.72
Distance to closest intersection (km) ³	0.73	0.40	0.88	0.51	0.78	0.38
Distance to the White House (km) ⁴	10.58	0.87	16.60	5.73	15.71	2.29
Pct of area in agricultural use or forest (adjusted for protected lands) in 1973 ⁵	0.00	0.00	0.12	0.25	0.05	0.10
Pct of dwelling units built post 1960 in 1970 ⁶	0.72	0.35	0.62	0.44	0.68	0.39
Pct of dwelling units in apartments 1970 ⁶	0.41	0.42	0.11	0.28	0.10	0.24
Pct zoned residential high density 1961 ⁷	0.16	0.28	0.04	0.14	0.02	0.10
Pct zoned residential low density 1961 ⁷	0.55	0.38	0.74	0.35	0.86	0.25
Employment density in 1972 (workers/ha) ⁸	37.76	44.97	19.96	27.75	9.71	12.73
Observations	32		161		61	

Included in the station areas are those TAZs that have their centroid within a mile of a station. Station areas that opened in 1978 are Takoma and Silver Spring. The station areas that opened in 1984 are Bethesda, Medical Center, Grosvenor, White Flint, Twinbrook, Rockville, and Shady Grove. The remaining three stations in Montgomery, Forest Glen, Wheaton and Glenmont, opened in the 1990s.

¹ Based on tract level averages from the 1970 Census of Population

² Based on tract level averages from the 1970 Census of Population re-distributed using dwelling units information from the Maryland PropertyView database

³ Distance from the centroid of the tract to the closest major intersection as determined by 1971 Maryland highway maps

⁴ Distance from the centroid of the tract to the White House

⁵ Based on a Maryland Department of Planning land use map for 1973

⁶ Based on information extracted from the Maryland PropertyView 1999 and 2002 databases

⁷ Based on a MNCPPC (1961) zoning map

⁸ Based on Green and James (1993) employment data, redistributed to the TAZs using land use data

4.2 Propensity scores

As described in Chapter 3, the variables used to describe the decision of where to open a Metro station include information on the location of the TAZ within the metropolitan area, on population, on employment, on past growth, on accessibility, and on historic land use. These aspects are also the factors that we generally think determine subsequent development patterns. The variables are used to calculate the probability of a TAZ (centroid) falling within a mile of a Metro station, or, in other words, being in the treatment group ($T=1$). Using the standard normal distribution, the equation to be estimated is

$$P(T = 1|X) = \int_{-\infty}^{\beta'X} \phi(z) dz \quad (4.1)$$

where X includes the above-mentioned initial conditions thought to influence the decision of where to locate a Metro station. The probability of a TAZ being located within the treatment distance is slightly different in 1990 and 2000, since additional TAZs were delineated based on the 2000 Census and some boundaries were slightly moved.

Table 4.2 shows the results from the propensity score estimations for the two sets (2000 and 1990) of TAZ boundaries. The hypothesis that the coefficients are jointly equal to zero is rejected for both 2000 and 1990. (The $\chi^2(12)$ has a value of 278 and 269, respectively.)

Table 4.2: Propensity Score Calculation for Montgomery County

Variable	Year 2000		Year 1990	
	Coef.	Std. Err.	Coef.	Std. Err.
Pct of population white in 1970 ¹	-0.959	0.943	-0.952	0.921
Mean household income in 1969, (\$1000) ¹	0.057	0.023 **	0.039	0.023 *
Mean income (\$1000) squared in 1969	-0.004	0.001 ***	-0.004	0.001 ***
Population density in 1970 (people/ha) ²	0.007	0.003 *	0.007	0.003 *
Distance to closest intersection (km) ³	-0.540	0.103 ***	-0.483	0.099 ***
Distance to the White House (km) ⁴	-0.095	0.012 ***	-0.098	0.012 ***
Pct of area in agricultural use or forest (adjusted for protected lands) in 1973 ⁵	-0.308	0.255	-0.255	0.247
Pct of dwelling units built post 1960 in 1970 ⁶	0.007	0.164	0.060	0.165
Pct of dwelling units in apartments 1970 ⁶	-0.896	0.331 ***	-0.628	0.316 **
Pct zoned residential high density 1961 ⁷	-2.088	0.541 ***	-1.960	0.549 ***
Pct zoned residential low density 1961 ⁷	-1.166	0.345 ***	-1.062	0.340 ***
Employment density in 1972 (workers/ha) ⁸	0.019	0.005 ***	0.015	0.004 ***
Constant	4.001	1.017	4.094	0.984
Number of observations	1252		1007	
coeff=0, Chi2 (12)	278	***	269	***

* significant at 10%; ** significant at 5%; *** significant at 1%

The propensity score is used to calculate the probability of a TAZ (centroid) of being within a mile of a Metro station. Given that the TAZ boundaries change slightly between 1990 and 2000, a separate score is calculate for each each.

¹ Based on tract level averages from the 1970 Census of Population

² Based on tract level averages from the 1970 Census of Population re-distributed using dwelling units information from the Maryland PropertyView database

³ Distance from the centroid of the tract to the closest major intersection as determined by 1971 Maryland highway maps

⁴ Distance from the centroid of the tract to the White House

⁵ Based on a Maryland Department of Planning land use map for 1973

⁶ Based on information extracted from the Maryland PropertyView 1999 and 2002 databases

⁷ Based on a MNCPPC (1961) zoning map

⁸ Based on Green and James (1993) employment data, redistributed to the TAZs using land use data

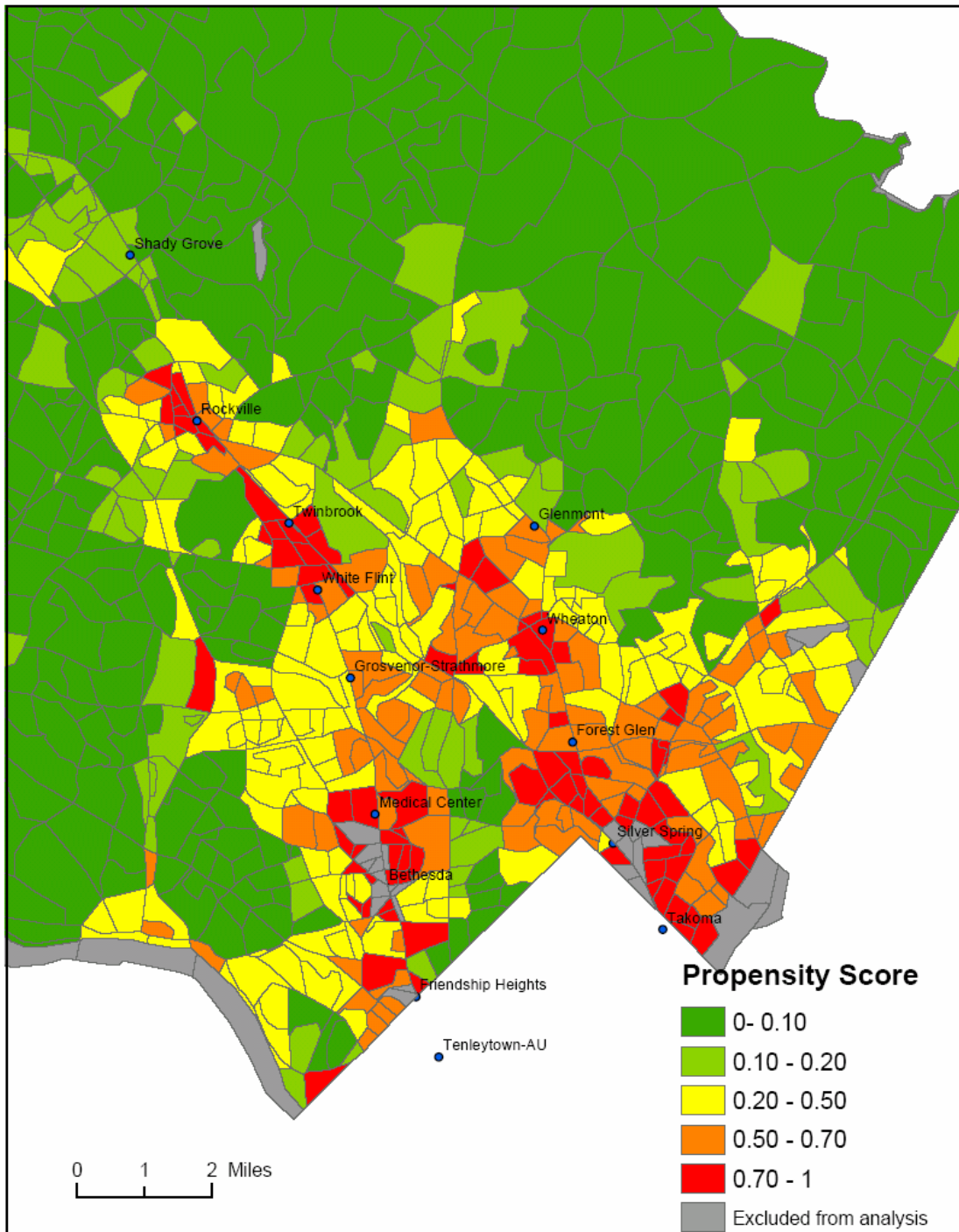
The coefficient estimates are similar for the two sets of TAZs, allowing for comparisons of the outcome impacts in the two post-Metro time periods. The correlation coefficient for the predicted probabilities for those TAZs that exist in both years is 0.98. Most of the variables included are statistically significant at the 95% level of confidence and in general have the expected signs. *A priori*, we would expect stations to be sited in the middle-income neighborhoods since purchasing potential and usage probability by the residents was one of the considerations. Income has the expected quadratic relationship with the probability of a TAZ being within a mile of a Metro station. This finding concurs with mode choice studies that find that the wealthiest and poorest households are less likely to use public transport. The population density measure is statistically significant at the 10 percent level of significance. The importance of population density is also expected since the network was built to connect the suburban population with the District employment. Given the relatively low variability in the percentage of minorities at the tract level, it is not surprising that the percentage of the population that is white is not significant in explaining the location of the Metro stations. Although the siting process did not explicitly attempt to join the suburban employment centers with the District, employment density, after controlling for all the other factors, does explain Metro station locations and is in line with the general planning vision presented in the 1961 general land use plan.

It is surprising that both the share of dwelling units in apartments and the percentage of land area zoned for high-density residential development are *negatively* correlated with the probability of a Metro station. Part of the reason for these negative

correlations may be that income is not accurately measured and thus the measures proxy, at a more local level, for lower income neighborhoods. Although neither the percentage of area devoted to agricultural uses and forestry nor the percentage of dwelling units built in the 1960s are statistically significant in explaining the location of a Metro station, they are arguably fundamental in capturing the development potential and the “recent” development patterns which are important in determining future development. The results of the propensity score further emphasize the fact that indeed the TAZs near future Metro stations were quite different from those TAZs that were further away.

Map 4.7 spatially depicts the distribution of the propensity scores for the 2000 boundaries. Many of the control TAZs with higher propensity scores are adjacent to the treatment areas. There are, however, some control TAZs with high propensity scores along the County’s southeastern border. Given the close proximity of the control and treatment impacts, if the treatment radius does not truly cover the area within the influence of the transit stop, then spillovers are possible. If so, the treatment impacts are a lower bound on the true impacts.

Map 4.7 Propensity Score for Montgomery County



Propensity scores calculated using the 2000 TAZ boundaries.

Table 4.3 and Figure 4.1 show the distribution of the propensity scores for treatment and control TAZs for both 1990 and 2000 boundaries. It is clear that there are few TAZs in the treatment group with low propensity scores and few TAZs in the control group with high propensity scores. In order to see what the treatment impacts are without these low density tails of the distributions, I follow a methodology adopted by Black and Smith (2004) of determining the treatment impact for observations in a “thick support” region. A thick support region is a range of propensity scores where “there are a substantial number of observations in both the treatment and comparison groups” (Black and Smith, 2004, pg. 118). Given the distribution of propensity scores in this study, I choose a thick support region that spans the propensity score between 0.2 and 0.7. This range avoids the use of single control observations excessively to construct the counterfactual, and generates sufficient observations for analysis. There are, in this case, 130 treatment observations and 151 controls. When the thick support is defined as those TAZs with a propensity scores between 0.3 and 0.7, or between 0.2 and 0.6, there results remain similar. I pinpoint any significant differences in Section 4.4.

Table 4.3a: Propensity score distribution of TAZs in Montgomery County, 1990 TAZ

Propensity Score range		Number of observations	
Max	Min	Control	Treatment
1	0.98	0	15
0.98	0.9	3	31
0.9	0.8	3	23
0.8	0.7	9	27
0.7	0.6	28	29
0.6	0.5	28	35
0.5	0.4	24	33
0.4	0.3	32	18
0.3	0.2	41	24
0.2	0.1	83	11
0.1	0	497	12

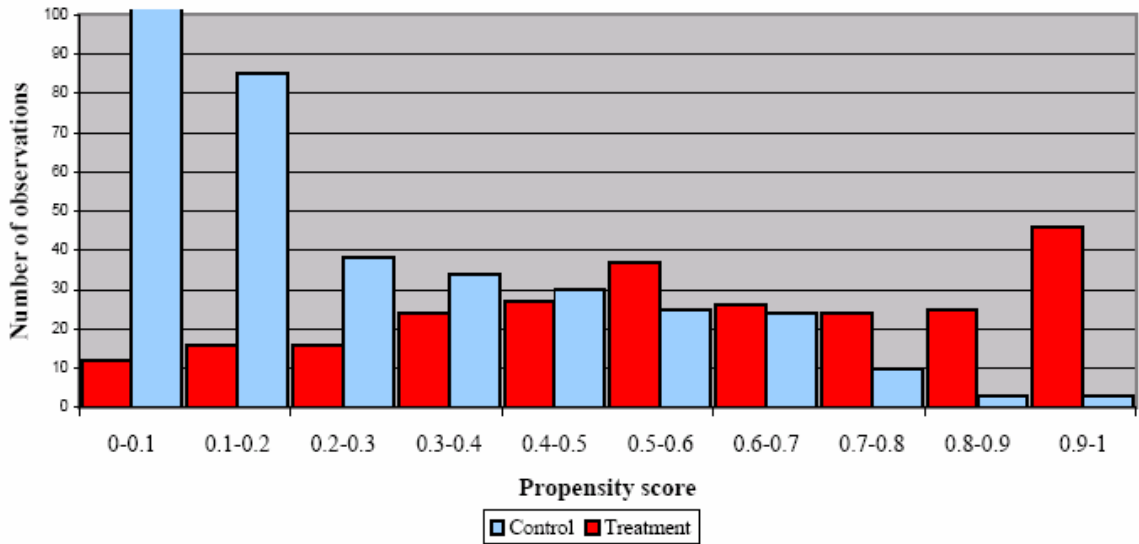
Highlighted rows make up the thick support region for the outcome analyses.

Table 4.3b: Propensity score distribution of TAZs in Montgomery County, 2000 TAZ

Propensity Score range		Number of observations	
Max	Min	Control	Treatment
1	0.987		20
0.987	0.9	3	26
0.9	0.8	3	25
0.8	0.7	10	24
0.7	0.6	24	26
0.6	0.5	25	37
0.5	0.4	30	27
0.4	0.3	34	24
0.3	0.2	38	16
0.2	0.1	85	16
0.1	0	747	12

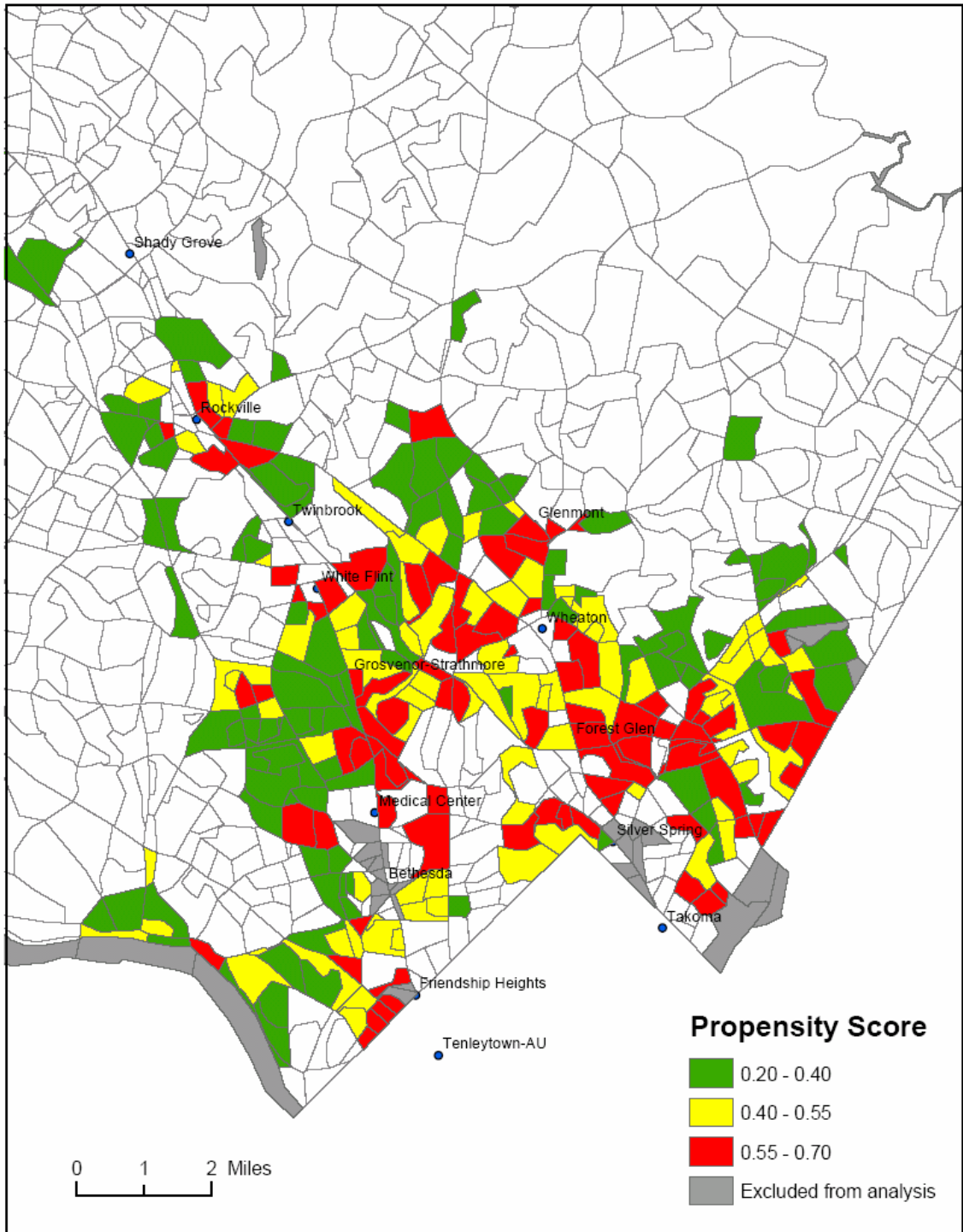
Highlighted rows make up the thick support region for the outcome analyses.

Figure 4.1: Distribution of propensity scores, Montgomery County 2000



The impacts calculated using the thick support sample apply only for the treated with moderate probabilities of receiving the treatment. The majority of the dropped treatment TAZs have propensity scores greater than 0.7, although some low propensity score observations are also dropped. In the thick support analyses there are no TAZs, for example, near the Shady Grove station. This area was unlikely (given the characteristics in the 1970s) to receive a station since it was much further away and had low levels of development. Map 4.8 graphically depicts the TAZs included in the thick support analyses.

Map 4.8 Thick support TAZs for Montgomery County



Propensity scores calculated using the 2000 TAZ boundaries.

4.3 Construction of the counterfactual

In order to estimate the counterfactual outcomes ($y^0 / T=1$), I use the Epanechnikov kernel, given its relatively good small sample properties (Frölich, 2004a).⁵⁵ That is, in Equation (4.2) the weight matrix, $W(i,j)$ for the outcomes in TAZs farther than a mile from a station is given by

$$W(i, j) = \frac{K\left(\frac{\hat{P}_j - \hat{P}_i}{h}\right)}{\sum_{k \in I_0} K\left(\frac{\hat{P}_k - \hat{P}_i}{h}\right)} \quad (4.2)$$

where h is the bandwidth of the kernel, and \hat{P}_i and \hat{P}_j are the probabilities of receiving treatment for a treatment observation i and a control observation j , respectively. I_0 is the set of possible control observations and $K(z) = 0.75(1 - 0.2z^2)/\sqrt{5}$ if $|z| < \sqrt{5}$ and is 0 otherwise. The Epanechnikov kernel has the desirable characteristic of placing greater weight on control observations with propensity scores more similar to the treatment observation and less weight on observations with more dissimilar propensity scores. It imposes automatically the common support condition discussed in Chapter 2 (Black and Smith, 2004). Besides using the Epanechnikov kernel, I also drop those treatment observations i for which $\hat{P}_i > \max(\hat{P}_j) \forall j \in I_0$. That is, treatment observations that had a higher probability of being within a mile of a station than any non-treated observation are not included in the average treatment impact.⁵⁶

⁵⁵ Black and Smith (2004) also find the Epanechnikov kernel to be superior to the Gaussian kernel.

⁵⁶ The Epanechnikov bases the common support on an interval of values such that it is possible for a treatment TAZ with a high propensity score to have the counterfactual built from control observations with only lower propensity score values. This excludes 20 treatment TAZs with a propensity score higher than 0.987 for 2000 impacts and 15 observations with a propensity score higher than 0.98 for 1990 impacts.

As discussed in Chapter 2, the choice of bandwidth is critical since it determines the range of propensity scores from which the counterfactual is built. When the bandwidth is larger, more control observations are used to determine the counterfactual. However, this implies that more weight is given to the observations that are more dissimilar from the treatment observation. A very small bandwidth matches the treatment observation to more similar controls, but less information is used and it is possible that if the bandwidth is small enough there are no control observations within the specified distance for a particular treatment observation. I follow the methodology adopted by Black and Smith (2004) and Frölich (2004a, 2004b) and determine the bandwidths using cross-validation outlined in Chapter 2. A different bandwidth needs to be determined for each of the outcome measures described in Chapter 3 and for each of the two outcome years. Eight different bandwidths are tested for each outcome measure based on the information from the non-station area TAZs.⁵⁷ The bandwidth with the lowest mean squared error is chosen as the cross-validated bandwidth. Table 4.4 gives the bandwidths used for each of the outcome measures. They range from 0.1 to 0.4 and are similar to those used in the literature.⁵⁸

⁵⁷ The bandwidths tested are 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, and 0.6.

⁵⁸ Black and Smith (2004) use bandwidths from 0.2 to 0.4 depending on the sample (men or women) and variables used for the propensity score. Frölich (2004a, 2004b) does not empirically apply the methodology.

Table 4.4: Optimal Bandwidth for Epanechnikov kernel for Montgomery County

Outcome measure	Outcome year	Cross-validated bandwidth
Employment density (workers/ha)	1990	0.4
	2000	0.1
Population density (people/ha)	1990	0.15
	2000	0.1
Development density (jobs+dwelling units/ha)	1990	0.15
	2000	0.1
Dwelling density using Census data (units/ha)	1990	0.1
	2000	0.1
Dwelling density using PropertyView data (units/ha)	1990	0.15
	2000	0.1
Dwelling density change 1970 to 2000 (units/ha)	1990	0.25
	2000	0.1
Dwelling density change 1990 to 2000 (units/ha)	1990	0.25
	2000	0.3
Percentage minority (%)	1990	0.15
	2000	0.2
Mean Household Income* (dollars)	1990	0.15
	2000	0.2

Optimal bandwidths are calculated using cross-validation, or leave-one-out, methods. The algorithm used for cross-validation was provided by Black and Smith which they used in Black and Smith (2004). The possible bandwidth tested were 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, and 0.6. The mean squared error did not vary greatly in this range.

Using the weight matrix calculated using the Epanechnikov kernel with cross-validated bandwidths, it is possible to determine the “weighted” initial conditions in the control group. That is, each control observation is weighted by the total weight it has in the construction of the average counterfactual and this weight is applied to the vector of

initial conditions. In order for the treatment and control group to be balanced in the vector of initial conditions (the x 's), the weighted means of the covariates cannot be statistically different across the two samples. The optimal bandwidth is important in determining whether or not the samples are balanced. When the bandwidth is larger, control TAZs with more dissimilar propensity scores are given relatively more weight than when the bandwidth is smaller. Thus, in general, the differences in the average initial conditions between the two groups become larger with larger bandwidths. For these analyses the samples are balanced (in this statistical sense) unless otherwise noted.

Tables 4.5 and 4.6 summarize the initial conditions—for both the common support sample and the thick support sample—of the treatment and control TAZs using the 2000 and 1990 boundaries, respectively. The tables report both the unweighted as well as the weighted means of the variables used in the matching. The weighted means are based on the Epanechnikov kernel with a bandwidth of 0.1. The unweighted sample includes all the observations in the sample. The weighted sample size for the treated observations is smaller since treatment observations i for which $\hat{P}_i > \max(\hat{P}_j) \forall j \in I_0$ are excluded from the sample.⁵⁹ The tables also give the probability that the two sample means are equal.

⁵⁹ By not imposing the condition, the results remain similar in terms of their statistical significance. The impacts on employment density and dwelling unit density are larger.

Table 4.5: Comparison of initial conditions in unweighted and weighted samples, Montgomery 2000 TAZs

Variable	Sample	All observations			Thick support		
		Mean		p> t ¹	Mean		p> t ¹
		Treated	Control		Treated	Control	
Pct of population white in 1970 (%)	Unweighted	0.93	0.91	0.18	0.95	0.96	0.33
	Weighted	0.93	0.95	0.16	0.95	0.96	0.42
Mean household income in 1969 (\$1000)	Unweighted	13.70	14.52	0.39	16.06	15.49	0.56
	Weighted	14.07	13.99	0.58	16.06	15.02	0.35
Mean household income (in \$1000) squared	Unweighted	230.90	284.21	0.20	276.33	284.66	0.73
	Weighted	240.97	241.47	0.47	276.33	266.33	0.66
Population density in 1970 (people/ha)	Unweighted	27.31	8.20	0.05	26.25	27.21	0.79
	Weighted	25.23	26.48	0.72	26.25	28.89	0.53
Distance to closest intersection (km)	Unweighted	0.84	1.77	0.05	0.90	0.94	0.59
	Weighted	0.87	0.91	0.26	0.90	0.88	0.73
Distance to the White House (km)	Unweighted	15.60	28.06	0.04	15.60	15.55	0.94
	Weighted	16.02	15.83	0.59	15.60	15.01	0.41
Pct of area in agricultural use or forest (adjusted for protected lands) in 1973	Unweighted	0.09	0.53	0.04	0.07	0.12	0.26
	Weighted	0.10	0.12	0.24	0.07	0.11	0.36
Pct of dwelling units built post 1960 in 1970	Unweighted	0.65	0.47	0.10	0.74	0.68	0.39
	Weighted	0.66	0.70	0.28	0.74	0.69	0.48
Pct of dwelling units in apartments 1970	Unweighted	0.14	0.04	0.09	0.08	0.11	0.55
	Weighted	0.12	0.13	0.48	0.08	0.13	0.36
Pct zoned residential high density 1961	Unweighted	0.05	0.01	0.10	0.05	0.05	0.96
	Weighted	0.05	0.04	0.57	0.05	0.06	0.75
Pct zoned residential low density 1961	Unweighted	0.74	0.96	0.04	0.89	0.88	0.84
	Weighted	0.78	0.83	0.13	0.89	0.86	0.49
Employment density in 1972 (workers/ha)	Unweighted	19.82	2.21	0.04	7.05	5.10	0.30
	Weighted	14.07	13.29	0.14	7.05	5.49	0.37
Number of observations	Unweighted	253	999		130	151	
	Weighted	233	999		130	151	

Weights are based on an Epanechnikov kernel with a 0.1 bandwidth. The thick support propensity scores range from 0.2 to 0.7.

Calculated using *pstest* algorithm in Stata written by Edwin Leuven and Barbara Sianesi.

¹ Gives the probability of the equality of the treated and the controls using Hotelling's T-squared test for each propensity score quantile.

Table 4.6: Comparison of initial conditions in unweighted and weighted samples, Montgomery 1990 TAZs

Variable	Sample	All observations			Thick support		
		Mean		p> t	Mean		p> t
		Treated	Control		Treated	Control	
Pct of population white in 1970 (%)	Unweighted	0.93	0.94	0.55	0.95	0.96	0.42
	Weighted	0.94	0.95	0.15	0.95	0.96	0.50
Mean household income in 1969, (\$1000)	Unweighted	13.46	17.03	0.10	15.92	15.29	0.54
	Weighted	13.85	13.78	0.58	15.92	14.99	0.41
Mean household income (in \$1000) squared	Unweighted	225.76	352.36	0.09	279.33	277.06	0.92
	Weighted	235.16	237.55	0.47	279.33	267.51	0.62
Population density in 1970 (people/ha)	Unweighted	27.53	10.51	0.06	25.13	25.61	0.89
	Weighted	25.38	26.46	0.66	25.13	27.46	0.56
Distance to closest intersection (km)	Unweighted	0.85	1.49	0.06	0.91	0.96	0.53
	Weighted	0.88	0.93	0.26	0.91	0.92	0.86
Distance to the White House (km)	Unweighted	15.66	23.56	0.04	15.74	15.94	0.76
	Weighted	15.98	15.62	0.91	15.74	15.48	0.69
Pct of area in agricultural use or forest (adjusted for protected lands) in 1961	Unweighted	0.09	0.40	0.05	0.09	0.12	0.38
	Weighted	0.10	0.12	0.30	0.09	0.11	0.50
Pct of dwelling units built post 1960 in 1970	Unweighted	0.64	0.45	0.10	0.74	0.66	0.34
	Weighted	0.66	0.71	0.23	0.74	0.68	0.42
Pct of dwelling units in apartments 1970	Unweighted	0.16	0.05	0.10	0.10	0.11	0.75
	Weighted	0.14	0.13	0.40	0.10	0.13	0.51
Pct zoned residential high density 1961	Unweighted	0.05	0.01	0.12	0.04	0.04	0.92
	Weighted	0.05	0.05	0.70	0.04	0.05	0.75
Pct zoned residential low density 1961	Unweighted	0.75	0.95	0.05	0.88	0.89	0.83
	Weighted	0.78	0.82	0.16	0.88	0.87	0.78
Employment density in 1972 (workers/ha)	Unweighted	19.83	2.99	0.05	6.47	5.41	0.50
	Weighted	14.37	13.29	0.17	6.47	6.06	0.77
Number of observations	Unweighted	259	748		125	153	
	Weighted	243	748		125	153	

Weights are based on an Epanechnikov kernel with a 0.1 bandwidth. The thick support propensity scores range from 0.2 to 0.7.

Calculated using *pstest* algorithm in Stata written by Edwin Leuven and Barbara Sianesi.

¹ Gives the probability of the equality of the treated and the controls using Hotelling's T-squared test for each propensity score quantile.

As expected, when the *unweighted* means for the treated and non-treated (based on all the observations) are compared, the Metro stations were located in areas that were more developed (both in terms of housing and employment), with a greater accessibility, closer to the White House and with more growth in the 1960s than an average TAZ. In many cases the null hypothesis of equal means is rejected at 90 percent level of confidence.

By weighting the control observations and by removing those treatment observations i for which $\hat{P}_i > \max(\hat{P}_j) \forall j \in I_0$ the differences between the two groups decrease. After the matching the null hypothesis of equal means cannot be rejected at 90 percent level of confidence for any of the variables capturing the initial conditions. That is, we have two groups that are statistically similar in their initial conditions. It should be noted that in some aspects (employment density in 1972, and percent of the TAZ zoned low-density residential in 1961) the two groups are barely similar. Consequently, there are possible gains in sample similarity from limiting the sample to a certain “thick support” range of propensity scores.⁶⁰

The second set of columns in Tables 4.5 and 4.6 summarize the initial conditions for the thick support sample. This subset of TAZs is statistically similar in the initial conditions even before matching.⁶¹ Unlike when all the observations are used, now variables such as employment density and percentage of land zoned for low-density residential uses on average are similar. By weighting the control observations the differences are reduced even further. Both of these observations highlight the fact that the thick support sample is more homogenous in its characteristics than the common support sample. The cost of achieving a more homogenous sample is a reduction in the sample size and the loss of generality of the treatment impacts.

⁶⁰ If the propensity scores higher than 0.975 are not excluded from the treatment group, the two samples remain balanced. In fact, the differences between the average initial conditions become smaller.

⁶¹ That is, when each control observation is in effect given a weight of one.

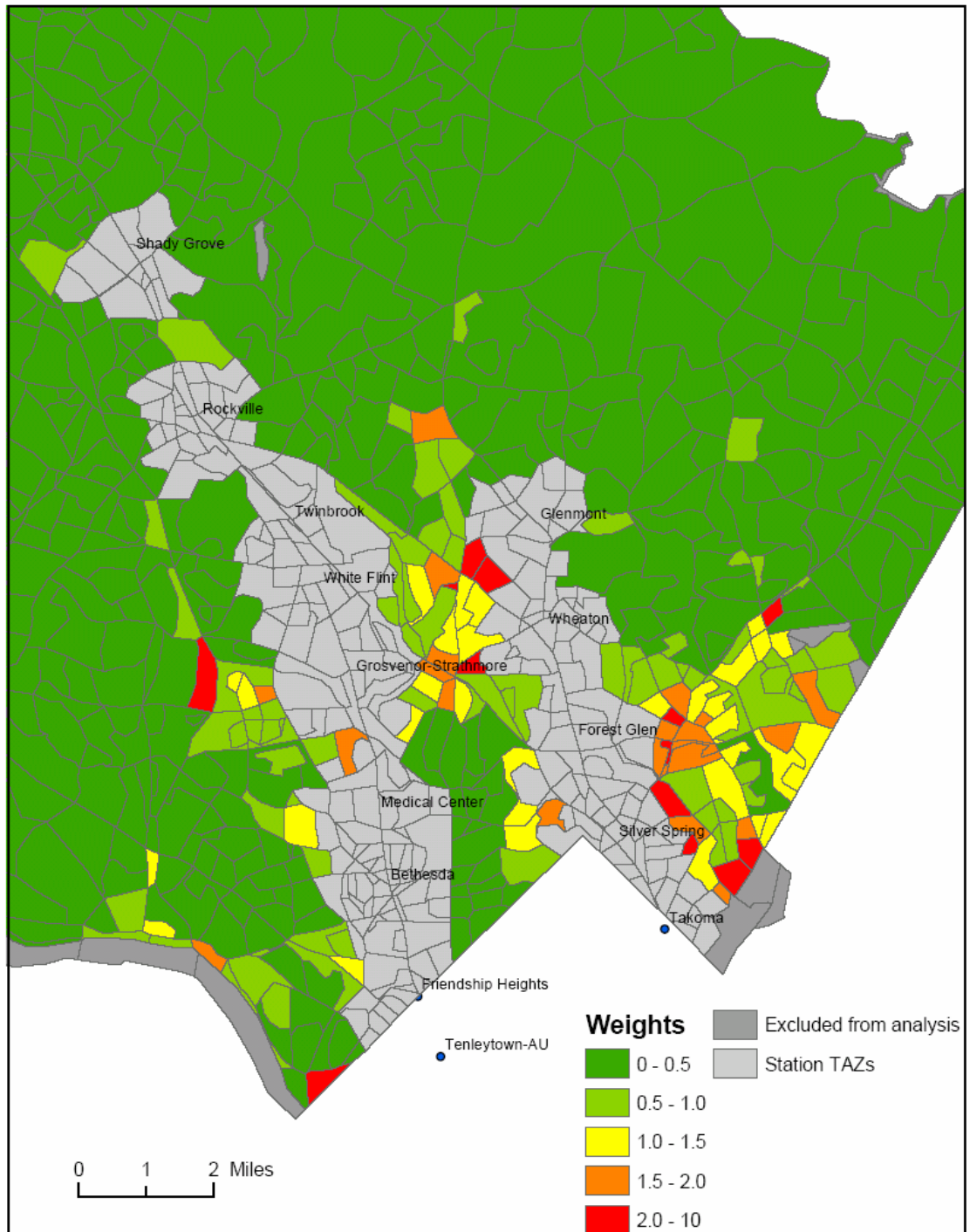
The controls with greater weight are in general within 19 miles of the White House and closely clustered. In the common support sample the weights range from 0.01 to 9.12 and there are 106 control TAZs with weights greater than one and 53 TAZs with weights greater than two. In the thick support sample the weights range from 0.21 to 1.51 and there are 51 control observations with a weight greater than one, but only 6 observations with a weight greater than two.⁶² The thick support sample relies less heavily on any individual control TAZs for the average counterfactual.

Map 4.9 shows an example of the weights used to construct the counterfactual in the outcome analyses. It is based on the Epanechnikov kernel with a bandwidth of 0.1 using the 2000 TAZ boundaries. Given that there are only two control observations that have a propensity score greater than 0.9 the matching algorithm relies heavily on these to construct the counterfactual. Most of the controls with weights greater than 0.5 are located close to the District line. Furthermore, many of the controls, especially those that have a greater weight, are adjacent to the treatment TAZs. This suggests the possibility of spillover effects if the treatment radius is misspecified.

⁶² This is based on the Epanechnikov kernel with a bandwidth of 0.1 and 2000 TAZ boundaries. The total number of control observations for the common support sample is 999 TAZs and for the thick support sample it is 151 TAZs.

Map 4.9

Weights for controls in Montgomery County, 2000



Propensity scores calculated using the 2000 TAZ boundaries.

Another observation from Tables 4.5 and 4.6 is that even though the TAZ boundaries changed somewhat between 1990 and 2000 and the area around Glenmont became part of the treatment in 2000 whereas in 1990 it was outside of the analyses, the average initial conditions remain relatively similar. This allows the comparison of the impacts from one year to another without having to attribute part of the impacts to differences in initial conditions.

4.4 Impacts of Metro stations in Montgomery County

The treatment impacts are calculated using the matching estimator given by

$$\hat{M}(T) = \sum_{i \in I_1 \subset CS} \omega(i) [y_i^1 - \sum_{j \in I_0} W(i, j) y_j^0] \quad (4.3)$$

where $\omega(i) = 1/N_1$ and N_1 is the number of treated observations and the W matrix is given by Equation 4.2. The outcome variables of interest are: employment, population and overall development densities as well as the income and racial makeup of the TAZs. Following I give the treatment impacts for 2000 and for 1990 for both the one-mile radius as well as the half mile radius. Also, I calculate the impacts for both the common support sample as well as the “thick support” sample described above.

Although the propensity score is calculated using all the TAZs not excluded by the conditions described in Section 3.2.3, in the treatment impacts only those station areas that were open at the evaluation date are included. For example, in the 1990 evaluation of Montgomery County, I exclude from the analysis (both from the treatment and control groups) those TAZs that are within a mile of the Glenmont station since the station opened in 1997.

4.4.1 *Impacts in 2000*

Using the common support sample, the 2000 outcomes show positive average treatment impacts from being within a mile of a Metro station for both the employment and population measures. In the thick support sample there are no statistical differences in the outcome measures for the 1-mile radius but the impacts are statistically significant for the half-mile radius treatment subsample.

4.4.1.1 One-mile treatment common support sample

The first set of columns in Table 4.7 presents the differences in the various outcome measures between all the treatment TAZs and the unweighted controls. If we were to assume that all TAZs were equally likely to be within a mile of a Metro stop (which is equivalent to assuming that siting is exogenous), treatment impacts in 2000 would be sizable. For example, those TAZs within a mile of station have approximately eight times the employment density of TAZs further away, and thrice the population and dwelling densities. Also, housing in the treatment TAZs has been developed at a much faster rate. The population living near the stations is more likely to be non-white and has a statistically significantly lower income than the rest of the County. All of these differences are statistically significant at the 99% level of confidence.

Table 4.7: Average treatment impacts for one mile radius in 2000 for Montgomery County

	No matching ¹			Full sample				
	Average treated	Average control	Unmatched impact	Average treated	Average weighted control	Treatment impact	90% C.I. Lower Bound	90% C.I. Upper Bound
Employment density (workers/ha)	41.63	5.25	36.38	27.60	10.73	16.87	10.37	29.43
Population density (people/ha)	40.86	13.91	26.95	39.84	26.21	13.63	3.98	43.86
Development density (jobs+dwelling units/ha)	56.82	9.49	47.33	41.16	19.18	21.98	12.25	38.03
Dwelling density using Census data (units/ha)	16.84	5.19	11.65	15.09	9.81	5.28	2.47	14.26
Dwelling density using PropertyView data (units/ha)	16.07	5.17	10.90	12.69	9.51	3.18	1.15	5.74
Dwelling density change 1970 to 2000 (units/ha)	7.88	3.05	4.83	5.41	2.00	3.41	1.85	6.24
Dwelling density change 1990 to 2000 ² (units/ha)	1.23	0.34	0.89	1.00	0.17	0.83	0.51	2.07
Percentage minority (%)	0.33	0.25	0.08	0.32	0.36	-0.04	-0.09	0.00
Mean Household Income* (dollars)	96,366	116,773	-20,407	98,362	91,659	6,703	-777	14,492
Number of observations	254	1011		233	999			
* Number of observations				207	901			

The treated observations are those TAZs that have their centroid within a mile of a station. All other TAZs are controls. The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 4.4 and using *psmatch2* algorithm in Stata. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval. Statistically significant impacts are in boldface.

¹ All of the "No matching" impacts are statistically significant at the 99% level of confidence.

² The two samples are not balanced in the employment density measure at the 90% level of confidence.

However, given the endogeneity of the decision, and the resulting differences in the average initial conditions in the treatment and potential control TAZs, the unmatched results do not reflect true impacts. When the counterfactual is built using treatment probabilities, the impacts become smaller in magnitude and at times statistically insignificant. The results for the one-mile radius are presented in Table 4.7 in the second set of columns. In terms of employment, the areas that are within a mile of a station have around 17 workers per hectare more than areas with similar characteristics but further away from the station. Had the stations not been built, and if the Metro system did not increase the overall County employment level, both groups would have had on average an employment density of 19 workers per hectare.⁶³ Given that the average density in the matched control in 2000 is around 11 workers per hectare, the difference is significant also in absolute terms although less than half of the impacts observed in the naïve unmatched case. The average treatment TAZ is 27 hectares (0.1 square miles) resulting in a net gain in employment of 459 employees, and in an average 1-mile treatment area the net gain is 13,736 additional employment opportunities. This impact is sizable given that the average employment around Montgomery Metro stations is 14,232 workers. It seems that the high propensity score / employment density TAZs are to some extent driving the result. This possibility will be analyzed in the thick support sample analysis.

⁶³ It must be emphasized that we cannot measure the impact the Metro has had in attracting new employment to the County, only what the distributional impacts on of the new employment have been. That is, we do not have a counterfactual for what would have been the employment densities without the Metro system.

There are also statistical differences between the treatment and control groups in the population, dwelling unit and the overall development density measures. There is approximately 50 percent more population and dwelling units in the treatment areas than in the controls with similar initial characteristics. These translate to a treatment impact of approximately 11,385 inhabitants within a mile radius from a station. It is also evident that the treatment TAZs experienced higher rates of residential construction both in the period 1970 to 2000 as well as in the period 1990 to 2000. Even though the growth rate decreased in the latter period (from 1970 to 2000 an average impact of 1.14 additional housing units per hectare per decade versus 0.83 additional housing units per hectare from 1990 to 2000) the differences are much more marked in the latter decade (differences of 174% and 488%, respectively between the treated and the controls). The slow down in construction reflects a countywide slow down in construction between 1990 and 2000. In Montgomery between 1970 and 1980 the housing stock grew by 33 percent; between 1980 and 1990 it grew by 37 percent; and between 1990 and 2000 it grew by only 12 percent.

In terms of overall development (as measured by the sum of employment density and dwelling unit density measures), the treatment TAZs are twice as dense as the control TAZs. This result is expected given that both employment and dwelling densities are higher in the treated areas than in the controls. The results emphasize that development, in general, has densified around the stations.

In terms of the socio-economic composition of population, there is no evidence of a negative impact on lower income households and minorities. The average income of a household within a mile of a station is \$6,703 higher, but this difference is not statistically significant.⁶⁴ The average percentage of minorities in the treatment group is 32% and in the counterfactual it is slightly higher, 36%, but again the difference is not significantly different from zero at the 90% level of confidence. That is, given the population living in Montgomery County less wealthy households are as likely to be living within a mile of a station as wealthier households, implying no significant negative impacts on less wealthy families from potentially higher housing prices. These results are different from the naïve case where the households in the two groups differ in income and race. Furthermore, these results need to be interpreted cautiously given that in the propensity score calculation the racial make-up of the TAZs and average household income are measured at the much more aggregate tract level possibly masking some of the initial differences between TAZs.⁶⁵

4.4.1.2 One-mile treatment thick support sample

The significant treatment impacts found above are not, however, robust to the exclusion of those TAZs with propensity scores greater than 0.7 and less than 0.2 (Table 4.8). There is, in general, no evidence of a positive treatment impact within the one-mile treatment radius when the thick support sample is used. For example, the impact on employment density reduces to 1.5 additional workers per hectare and it is not

⁶⁴ The number of observations is slightly smaller as not all TAZs have population and thus no income information. These TAZs remain outside of the impact calculations.

⁶⁵ It is possible that even though at the aggregated level these variables were balanced in the initially, with more disaggregated data there may be differences. However, no data exist to test the possibility.

statistically significant at the 90 percent level of significance. Similarly, there are no statistically significant impacts on population, dwelling, and overall development densities. These results suggest that areas with the highest propensity scores (and greater initial densities) have in general a greater degree of post-treatment development as well. The majority of the treatment impact observed in the common support sample occurred in the TAZs with propensity scores greater than 0.7.

Table 4.8: Average treatment impacts for thick support TAZs in 2000 for Montgomery County

	Thick support, treatment one mile radius				Thick support, treatment half a mile					
	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound
Employment density (workers/ha)	12.71	11.20	1.51	-4.20	6.57	23.78	8.50	15.28	6.01	29.35
Population density (people/ha)	35.24	27.57	7.67	-1.88	24.60	59.55	28.97	30.58	-0.62	83.05
Development density (jobs-dwelling units/ha)	23.10	19.99	3.11	-2.78	8.99	37.86	17.14	20.72	10.32	36.09
Dwelling density using Census data (units/ha)	12.21	10.50	1.71	-0.99	4.67	15.62	10.83	4.79	-0.28	10.65
Dwelling density using PropertyView data (units/ha)	11.72	10.81	0.91	-1.78	3.93	18.54	11.48	7.06	0.54	16.86
Dwelling density change 1970 to 2000 (units/ha)	4.30	2.41	1.89	-0.23	4.62	9.91	2.73	7.18	1.00	15.11
Dwelling density change 1990 to 2000 (units/ha)	0.57	0.19	0.38	0.06	1.00	1.53	0.16	1.37	0.15	3.80
Percentage minority (%)	0.34	0.36	-0.02	-0.07	0.04	0.40	0.39	0.01	-0.07	0.09
Mean Household Income * (dollars)	105,743	91,953	13,790	1,785	25,955	81,559	87,128	-5,569	-19,702	10,131
Number of observations	130	151				34	133			
* Number of observations	128	134				33	124			

The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 4.4 and using *psmatch2* algorithm in Stata. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval. Statistically significant impacts are in boldface.

There is some evidence in the thick support sample of a positive impact on recent housing unit construction. Between 1990 and 2000 on average the impact from being in the treatment was 0.38 additional dwelling units per hectare. This impact is relatively small and additional time would be necessary to determine if there are any cumulative impacts in the long run. The only other significant difference is the mean household income. The income of households within a mile of a station is on average \$14,000 more than in the TAZs used to construct the counterfactual. The fact that there is a positive income difference signals a possible inability of less wealthy households from acquiring housing near Metro stations. Overall these results suggest that within the more comparable treatment TAZs within a mile of a Metro station, the stations have not affected the development patterns although there is some evidence of differences in the type of households.

When different ranges of propensity scores are considered as part of the “thick support region” the results do not change qualitatively. When the propensity score range is limited to 0.3 to 0.7, the results are qualitatively similar to those generated by a thick support of 0.2 to 0.7. When the range is limited from 0.2 to 0.6, the results are again qualitatively similar except the impact on the dwelling density change from 1970 to 2000 becomes statistically significant and the impact on mean household income ceases to be statistically significant, although the point estimates are of similar magnitudes.

To test for spatial spillovers I limit the control observations to those TAZs that are more than 1.5 miles from the closest station. That is, TAZs in a buffer between 1 mile

and 1.5 miles from the closest station are excluded from both the treatment and the control groups. If there are important spillover effects, such that the Metro stations have also impacted nearby TAZs in the control group, then the impacts calculated without buffering will be smaller than the true impacts. For the case of Montgomery County, the results do not change with the implementation of a buffer. That is, the point estimates for the treatment impacts are qualitatively similar and the statistically significant impacts remain statistically significant. The results from the impact analyses with a buffer are presented in Table 4.9.

Table 4.9: Average treatment impacts for thick support TAZs in 2000 for Montgomery County (possible control TAZs are at least 1.5 miles from the closest station)

	Inick support, treatment one mile radius				Inick support, treatment half a mile					
	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound
Employment density (workers/ha)	12.86	12.64	0.22	-7.61	6.31	22.39	8.54	13.85	4.83	28.82
Population density (people/ha)	35.71	26.67	9.04	-1.81	25.48	61.07	28.82	32.25	-0.34	93.62
Development density (jobs+dwelling units/ha)	23.41	21.11	2.30	-6.29	9.09	36.73	17.16	19.57	8.96	34.13
Dwelling density using Census data (units/ha)	12.34	9.82	2.52	-1.06	5.97	15.93	10.26	5.67	0.05	11.86
Dwelling density using PropertyView data (units/ha)	11.85	9.87	1.98	-1.70	5.81	18.95	10.62	8.33	0.79	17.85
Dwelling density change 1970 to 2000 (units/ha)	4.37	1.38	2.99	0.89	5.47	10.05	1.63	8.42	2.53	17.13
Dwelling density change 1990 to 2000 (units/ha)	0.58	0.20	0.38	0.03	1.01	1.41	0.16	1.25	0.02	3.45
Percentage minority (%)	0.35	0.38	-0.03	-0.10	0.04	0.39	0.41	-0.02	-0.13	0.08
Mean Household Income * (dollars)	105,873	90,164	15,709	343	28,312	82,354	86,815	-4,461	-20,706	12,160
Number of observations:	127	90				33	78			
* Number of observations:	125	76				32	70			

The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 4.4 and using *match2* algorithm in Stata. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval. Statistically significant impacts are in boldface.

In general, results from the thick support sample are more robust than those from the common support sample. In the thick support sample the two sets of TAZs are more similar in their initial characteristics (that is, they are more balanced) than when the extremes of the probability distribution are included. For the rest of the analyses only the thick support sample results are presented.

4.4.1.3 Half-mile treatment thick support sample

I use a subsample of TAZs within half mile of a station to test the hypothesis that impacts differ depending on the distance (even among the TAZs within a treatment distance). In this sample, there are 34 treatment TAZs and 133 TAZ to construct the counterfactual from, and the two groups are balanced.⁶⁶ Comparing the thick support results for TAZs within a mile of a station with the thick support results when only TAZs within half a mile of a station are included in the treatment observations, there are differences in the treatment impacts. The second set of columns in Table 4.8 summarizes the treatment impacts for the half-mile TAZs in the thick support sample.⁶⁷ As expected the treatment impacts are greater than in the one-mile treatment thick support sample. Furthermore, development outcomes are statistically significant.

The impact of a Metro station on employment is around 15 additional opportunities versus the 1.5 (statistically insignificant) additional jobs available in the one-mile treatment area. This impact translates to approximately 432 jobs more in an

⁶⁶ The sample is small but similar to the Black and Smith (2004) thick support samples. They have 44 and 39 treatment observations for their analyses.

⁶⁷ The controls are still drawn from those TAZs that are further than one mile from the closest Metro station.

average TAZ within half mile of a station and with the characteristics of a TAZ with a propensity score between 0.2 and 0.7. This is equivalent to 3,050 additional jobs within half mile of a station. When all of the 12 Montgomery County stations are considered, this translates to 36,600 additional jobs within half a mile of a Metro station or about 9% of the total number of jobs in the County. These additional jobs may be either relocations within the county due to the Metro system (or employment that would have existed in the county even without the Metro system), or jobs that were attracted by the Metro system and without it would have located outside of the county.

The impact on population is 31 additional people per hectare (or 6,300 additional inhabitants within the half mile radius), which is approximately 7 additional dwelling units per hectare.⁶⁸ Although the impact on population density is large, it is statistical insignificance at the 90 percent level of significance. The impact on dwelling units corresponds to the difference in the number of new dwelling units built between 1970 and 2000. The construction of new dwelling units has occurred at a greater pace for the TAZs within half mile of a station compared to those within a mile. The rate of new housing construction slowed down between 1990 and 2000 as in the one-mile sample. Both, the population and the dwelling unit density impacts are big when compared to the counterfactual densities of 29 people per hectare and 12 dwelling units per hectare.

⁶⁸ It is not clear why the dwelling density measure derived from the Census data does not capture treatment impacts whereas the measure derived from the Maryland PropertyView data does. It must also be noted that the Census measure is only marginally not statistically significant at the 90% level of confidence, whereas the PropertyView measure is just statistically significant at the 90% level of confidence.

There is no difference in the socio-economic factors between the treatment and the control groups in the half-mile sample. Therefore, it appears that the higher income households are located between half a mile and a mile from a station. Potentially the higher income households locate so as to avoid any negative externalities from a station or from the higher concentration of commercial activity but in a location where they still are able to enjoy the added accessibility benefits provided by the Metro for commute trips. These differences may also be due to the supply of housing with different characteristics in the two rings.

Depending on the alternative definition used for the thick support, the results change. When the thick support is given by those TAZs with propensity scores between 0.3 and 0.7, all the point estimates and the impacts significant at the 90% of confidence remain qualitatively similar. However, the results are not robust to changing the thick support region to propensity scores between 0.2 and 0.6. In this sample, the impact on employment density halves but remains statistically significant. The impacts based on both dwelling unit density measures become statistically significant at the 90% level of confidence and the point estimates double. Also, the impact on the change in the dwelling unit density between 1970 and 2000 doubles.

There is no evidence of spillover impacts in the thick support half-mile sample. When the control observations are made up of those TAZs that are further than 1.5 miles from the closest Metro station the results remain qualitatively similar. The results are presented in Table 4.9.

4.4.1.4 One-mile multiple treatment analysis

The above results suggest that the treatment impacts are greater in areas closer to the Metro station than in areas still within a walking distance but farther. In order to further explore the impact of varying doses of treatment, I calculate the treatment impacts using the multiple treatment framework outlined by Lechner (1999) and summarized in Chapter 2. Given the small sample size, only very preliminary results can be obtained for the multiple treatment case. The results are consistent with those from the binary treatment analyses.

The treatment impact in the multiple treatment case is given by the nearest neighbor matching algorithm

$$\hat{M}(T) = \frac{1}{N_t} \sum_{i \in I_t, \in CS} [y_i^t - \sum_{j \in I_s} W(i, j) y_j^s] \quad (4.4)$$

where the $W(i, j)$ is 1 for observation j in treatment s that is the $\min(d(i, j) \forall j \in I_s)$, where $d(i, j)$ is the closeness of the two conditional probabilities $\hat{P}_k^{s|ts}(x)$ and $\hat{P}_k^{t|ts}(x)$ for $\forall k \in \{I_t, I_s\}$. I assign three different treatment dosages. In the first group are those TAZs with centroids within half a mile of a station (strongest dose), the middle ring includes TAZs with a centroid between half a mile and a mile from a station, and the outer ring (weakest dose) includes the TAZs with a centroid farther than a mile from the closest Metro station.

In order to obtain the treatment impact, I first calculate the three different conditional probabilities for each TAZ. Each TAZ is associated with a vector of

probabilities reflecting the probability of receiving each of the three different doses of treatment. To obtain the treatment impact of receiving treatment t instead of treatment s , the closeness of the probabilities associated with those TAZs in either one of the two treatments is determined by the Mahalanobis distance measure discussed in Chapter 2.⁶⁹ The matching is done with replacement such that a particular observation in the “control” group may be the counterfactual for more than one treatment observation. As pointed out by Lechner (1999) each group acts as a treated and a control group and therefore it is not the case that the no-treatment group, or control, is always larger as in a binary treatment case. The cut-off probabilities implied by the general common support condition for the multiple treatment analysis are given in Appendix 4.1.

Statistically significant impacts are observed for the most dissimilar TAZs (the inner ring versus the outer ring) (column 1 in Table 4.10). These results are akin to the half-mile binary results presented in Table 4.8 and, in general, the same outcome measures have statistically significant treatment impacts. The impact on employment density is around 30 additional jobs per hectare and there are an additional 6 dwelling units per hectare. Again, there are no impacts in population density or the socio-economic outcome measures.

⁶⁹ Unfortunately given the small sample size it is not possible to use a kernel matching algorithm.

Table 4.10: Treatment impacts with varying dose of treatment in 2000 for Montgomery County

	<0.5 mile vs >1 mile impact	<0.5 mile vs 0.5>x<1 mile impact	0.5>x<1 mile vs >1 mile impact
Employment density (workers/ha)	29.79 *	11.11	2.76
Population density (people/ha)	10.89	18.83	-2.4
Development density	32.35 *	13.72	3.1
Dwelling Density (units/ha)	3.47	1.67	-0.2
Dwelling Density using PropertyView (units/ha)	6.13 *	0.09	-0.31
Dwelling Density change 1970 to 1990 (PropertyView) (units/ha)	7.63 *	-1.49	0.12
Dwelling Density change 1980 to 1990 (PropertyView) (units/ha)	2.11 *	2.05 *	0.08
Percentage minority	-0.12	0.07	-0.11
Mean household income	2,562	-18,480 *	24,353 *
Number of obs closer to stations	60	60	126
Number of obs further from station	201	126	201

Treated group is the one that contains the set of TAZ nearer to the closer Metro station. That is, in the first and second columns the treated are those less than a half a mile from a station, in the third column the treated are those TAZs that are between half a mile and a mile from a station. For the analyses the general common support condition is imposed. The cut-off probabilities are given in Appendix 4.1.

* Statistically significant at the 90% level of confidence as determined by 2000 bootstrap repetitions using bias adjusted intervals

Even though some of the point estimates for the differences in the density outcome measures between the TAZs in the inner ring versus those in the middle ring are quite large (column 2 in Table 4.10), they are not statistically significant. The treatment impacts from receiving a moderate versus a low dose are, even in their point estimates, negligible (column 3 in Table 4.10). The results suggest that for the densification impacts have not yet reached beyond the immediate vicinity of the stations. It is also possible that the benefits from locating within the middle ring are too small for any densification to occur.

The only other notable result from the multi-treatment analyses, besides the fact that the impacts vary with the dissimilarity of doses, is the impact on the distribution of income. The multi-dose analysis re-enforces the fact that the high-income households

live between half a mile and a mile from a Metro station. There is a “negative” impact on household income (-\$18,480) from being within a half mile of a station when compared to the middle dose and a positive impact on income (\$24,353) from receiving the middle dose rather than the weakest dose. That is, the lower income households live in the inner or outer treatment rings and the wealthiest in the middle ring. Again, these differences may reflect the type of dwelling units available at each distance.

4.4.2 Impacts in 1990

Besides the spatial differences in the impacts, one may also expect differences depending on the length of time a station has been in operation. To this end, I compare the impacts in the 2000 and 1990 thick support samples for both the one-mile and half mile radii.

The impacts in 1990 were quite similar to the impacts in 2000 when the TAZs within a mile of a station in the thick support sample are considered (Tables 4.11 and 4.8, respectively). There is no impact on employment, population or dwelling unit density. There was slightly more new development between 1970 and 1990 and between 1980 and 1990 in the treatment TAZs than in the counterfactual. However, these differences had not translated into statistically significant difference in dwelling density levels. Unlike in 2000, there was no statistically significant impact on household income in 1990. The difference in the impact between 1990 and 2000 is statistically significant at the 95% level of confidence. The impact on income has increased over time.

Table 4.11: Average treatment impacts for thick support TAZs in 1990 for Montgomery County

	One mile radius, thick support					Half a mile radius, thick support				
	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound
Employment density (workers/ha)	13.50	12.29	1.21	-4.17	6.59	25.36	13.97	11.39	0.93	24.39
Population density (people/ha)	24.59	23.82	0.77	-3.39	4.48	31.65	24.96	6.69	-2.82	17.64
Development density (jobs+dwelling units/ha)	22.31	20.05	2.26	-2.74	7.85	38.90	22.60	16.30	3.09	32.50
Dwelling density using Census data (units/ha)	10.56	9.66	0.90	-1.19	3.11	15.60	10.22	5.38	-0.21	12.84
Dwelling density using PropertyView data (units/ha)	11.63	9.94	1.69	-1.19	4.97	19.28	10.59	8.69	-0.41	23.55
Dwelling density change 1970 to 1990 (units/ha)	3.50	2.17	1.33	0.15	5.40	11.98	2.16	9.82	1.34	22.61
Dwelling density change 1980 to 1990 (units/ha)	1.75	0.78	0.97	0.25	1.78	3.97	0.64	3.33	0.90	6.26
Percentage minority (%)	0.17	0.20	-0.03	-0.07	0.01	0.17	0.22	-0.05	-0.11	0.03
Mean Household Income* (\$ 1989)	70,532	67,083	3,449	-3,767	9,505	60,073	65,448	-5,375	-16,616	4,550
Number of observations	125	153				27	153			
* Number of observations	119	134				24	134			

The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 4.4 and using *psmatch2* algorithm in Stata. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval. Statistically significant impacts are in boldface.

As in 2000, there are some positive impacts on densities in the sample of treatment TAZs within half mile of a station. In 1990, the impact on employment density was slightly smaller than in 2000 (11 workers per hectare versus 15 workers per hectare) but the difference is not statistically significant. That is, even though the point estimates suggest a densification process, it is not possible to rule out no additional densification in employment between 1990 and 2000. Again, possibly with another decade of data, statistically significant differences could be observed. In the half-mile thick support samples the differences in the other outcome measures between 1990 and 2000 are not statistically different either.

4.4.3 Western branch of Red Line

To gain further insight into the role of time, I look at the outcome impacts in 1990 and 2000 using only the stations that opened in 1984 (the western branch of the Red Line).⁷⁰ In this way, I eliminate some of the noise introduced by stations opening in different years. I calculate the impacts for the thick support sample. There are statistically significant impacts on employment and development density in 2000 but not in 1990 (Table 4.12).⁷¹ Again, given the large confidence intervals, it is not possible to assert statistically that there has been densification and the results can only “suggest” that such a mechanism is operating. It is possible that in 1990 when these stations had been open for only six years, insufficient time had passed for statistically significant differences to emerge. Ten years later we observe a treatment impact of 6 workers per

⁷⁰ The eastern branch stations opened in 1978, 1990 and 1998 leading to only a few treatment TAZs for each year.

⁷¹ Both of the samples are balanced in the covariates used to calculate the propensity score.

hectare. This difference in employment also drives the total development measure to be statistically significant in 2000 but not in 1990.

In terms of the socio-economic composition of the population in the treatment versus control areas we observe two trends along the western branch of the Red Line. First, the minority population tends to be located away from the vicinity of Metro stations and the difference appears to have increase in time (from a 7 percentage point difference to a 10 percentage point difference). However, this difference is not statistically significant. It is important to note that even though we do not see differences in the location of minorities when the whole system is analyzed, we do observe negative impacts on the minority population near the stations along the western branch.⁷²

Furthermore, the households closer to the Metro stations tend to be wealthier than in the comparable control areas. Similarly, this inequality is increasing through time. Whereas in 1990, the households in the treatment areas earned approximately 1.09 times that of those in the control areas (and the difference was not statistically significant), in 2000 the difference had grown to 1.32 times. The increase in the difference in the average household income is statistically significant at the 90% level of confidence.

⁷² When all the stations in Montgomery County are analyzed, the average differences are negligible given the high concentration of minorities around the Takoma and Silver Spring stations as observed in the first column of Table 4.1. Given the small number of observations in the thick support region around these two stations, which both opened in 1978, I cannot do a separate historical analysis for these two stations.

Table 4.12: Average treatment impacts for one mile radius in 2000 and 1990 for Montgomery County for stations on the Western branch

	Impacts in 2000					Impacts in 1990				
	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C.I. Upper Bound	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C.I. Upper Bound
Employment density (workers/ha)	16.64	10.53	6.11	0.05	13.10	15.09	11.71	3.38	-1.99	9.12
Population density (people/ha)	37.56	26.94	10.62	-3.11	36.75	23.96	23.46	0.50	-3.79	5.06
Development density (jobs-dwelling units/ha)	26.79	19.20	7.59	0.63	15.06	23.62	19.24	4.38	-1.11	10.15
Dwelling density using Census data (units/ha)	12.47	10.30	2.17	-1.21	6.56	10.66	9.48	1.18	-1.24	3.92
Dwelling density using PropertyView data (units/ha)	11.70	10.51	1.19	-1.99	5.55	12.17	9.75	2.42	-0.95	7.14
Percentage minority (%)	0.25	0.35	-0.10	-0.17	-0.05	0.13	0.20	-0.07	-0.11	-0.02
Mean Household Income* (\$ 1999)	123,110	93,579	29,531	14,783	47,028	99,183	91,157	8,026	-1,417	12,781
Number of observations	81	151				92	153			
* Number of observations	80	134				87	134			

The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 4.4 and using *psmatch2* algorithm in Stata. Only observations within the thick support region (propensity scores between 0.2 and 0.7) are used in the analyses. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval. Statistically significant impacts are in boldface.

4.5 Conclusions

Overall, the analyses show that the manner in which the counterfactual of no-Metro station is chosen is important. Whilst comparing the areas surrounding the Metro stations with areas farther away one observes large differences in employment, population and dwelling unit densities, these results are based on the faulty assumption that Metro stations were exogenously placed. The decision of where in the County to site a Metro station was based on factors such as employment and population density, distance from the CBD, accessibility, and land use. These conditions make the Metro station areas very different from areas where no Metro was built. Incorporating the decision process into the analyses the results change, at times, drastically. It is also evident that concentrating on those treatment observations within a range of propensity scores where there is a significant number of similar control observations, the impacts are lower than when the extremes of the probability distribution are included.

The results suggest that the development impacts depend both on the distance of the treatment TAZs from the Metro station as well as on the length of time that the particular station has been in operation. In the thick support sample, there are no positive impacts when all stations within a mile are considered but there are positive impacts on employment and dwelling units when only TAZs within half a mile of a station are considered. This result is consistent with the results from the analysis with multiple doses. Also, I find that the impacts become more prominent when the stations have been in operation for longer. These results are clearer when I restrict the treatments to those TAZs that are along the western branch of the Red Line where all stations have been

open for the same length of time. There is also some evidence of a possible inability of less wealthy households and minorities from acquiring housing near the Metro stations and that the difference in the average household income has increased over time. The results remain qualitatively similar to different definitions of “thick support regions” and there is no evidence of spillover impacts.

Chapter 5

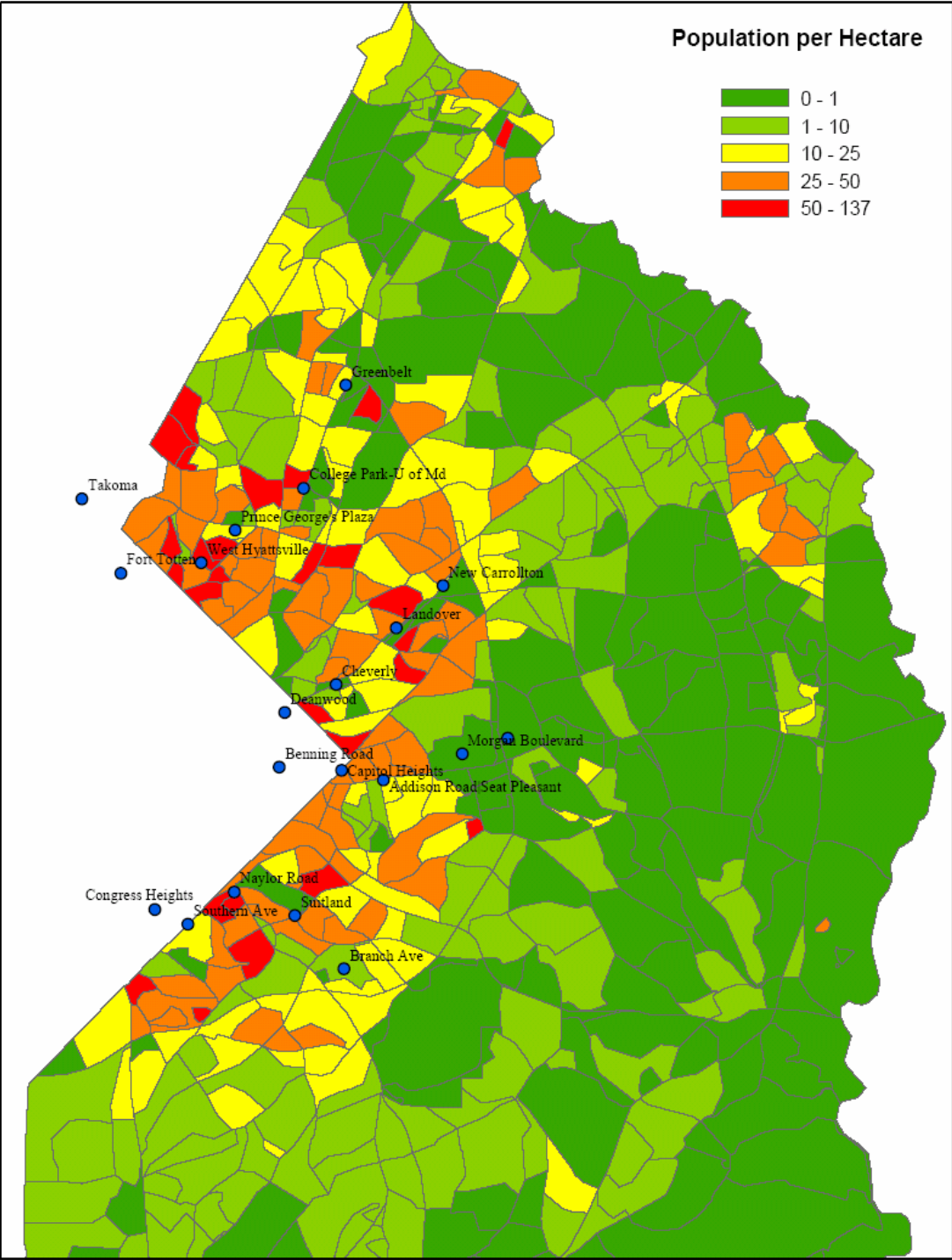
Case Study: Prince George's County

This chapter applies the same methodology used for Montgomery County to Prince George's County to determine the impacts of the Metro stations on development densities. First, I present an overview of the characteristics of the County in the early 1970s. Second, I analyze the propensity score using the 2000 boundaries. Third, I report the treatment impacts on employment, population, and the overall development densities obtained using the Epanechnikov kernel matching estimator. I compare and contrast the results between the two Maryland counties. Although there were some similarities in the development patterns between Prince George's County and Montgomery County prior to the opening of the Metro, there were also some significant differences possibly yielding different results.

5.1 Initial Conditions in Prince George's County

In 1970, the population in Prince George's County was mainly concentrated around the District border. There were also higher concentrations of population in Laurel, in the northeastern corner of the County, and in Bowie, in the eastern part of the County (Map 5.1). Although many of treatment TAZs had high population densities, there were TAZs within a mile of a station that had relatively low population densities. As in Montgomery County, the farther areas had population densities less than 10 people per hectare and, based on this measure alone, should provide poor controls for the majority of the treatment TAZs.

Map 5.1 Population Density in PG County, 1970

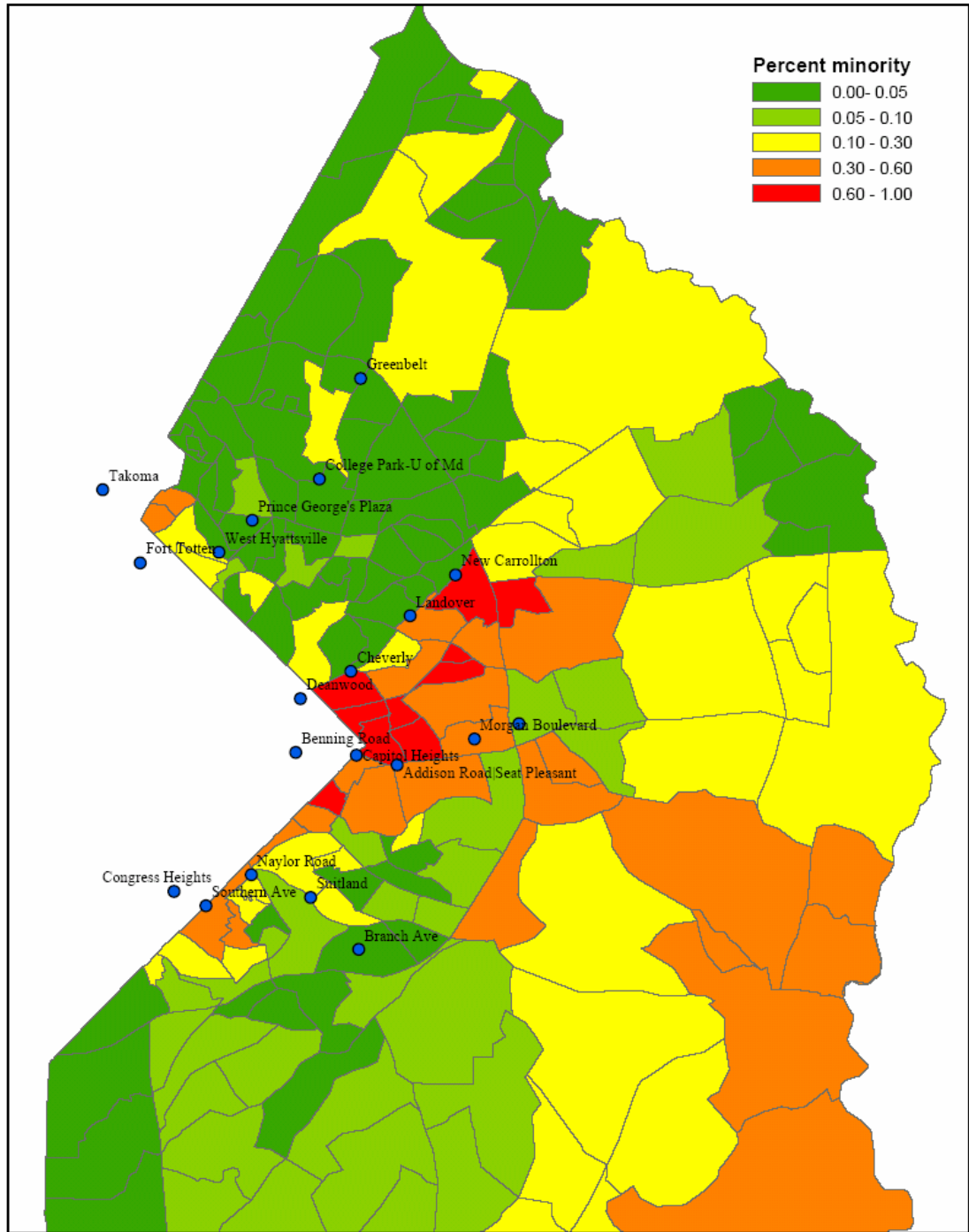


Based on the 1970 Census of Population; tract level data are redistributed to TAZs based on Maryland PropertyView

In the early 1970s, the socio-economic character of Prince George's County was quite different from Montgomery County. There was more racial diversity in Prince George's County than in Montgomery County. Whereas in Montgomery County there were few Census tracts with minority concentrations greater than 20%, in Prince George's County there were various tracts with a high proportion of minorities, including tracts (between Cheverly and Addison Road Metro stations) with over 90% of the population belonging to a minority (Map 5.2). The stations were placed in TAZs with high and low minority concentrations. For example, station areas around Greenbelt, College Park, Prince George's Plaza and West Hyattsville were predominantly white.

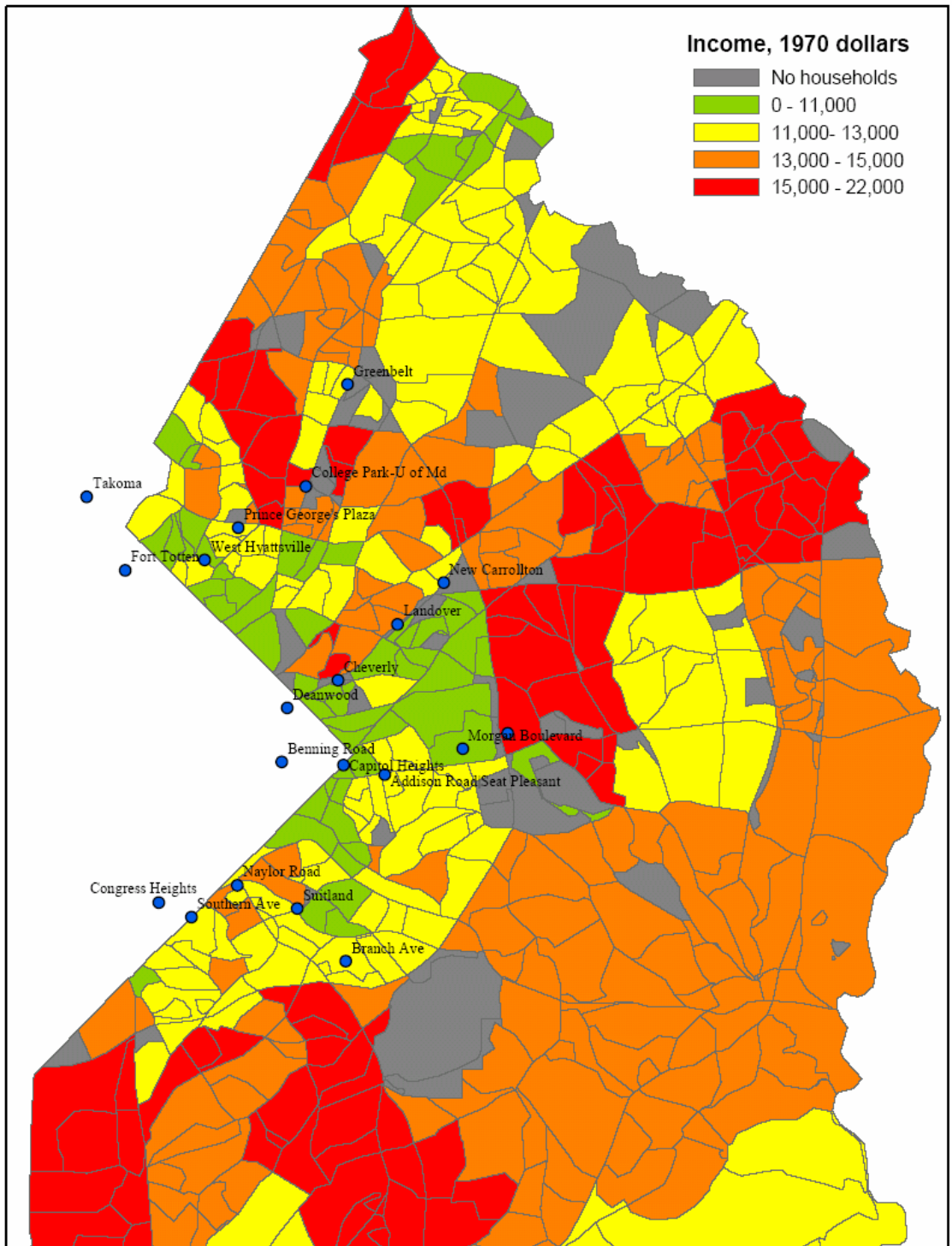
In terms of average household income, Prince George's County was much poorer than its neighbor. Unlike in Montgomery County, the majority of the poorer households lived close to the District border and not in the rural parts of the County (Map 5.3). Richer households of the County lived well beyond a mile from a planned Metro station. That is, the wealthiest were not targeted as the potential users. There was, however, a great variability in mean household incomes near the future Metro stations. For example, Metro stations such as Cheverly, Prince George's Plaza and College Park were to be located close to relatively affluent neighborhoods. Stations such as Deanwood and Capitol Heights were located in areas with the lowest average household incomes in the County. These low-income areas were also the areas with higher concentrations of minority populations.

Map 5.2 Percentage of Population Belonging to a Minority in Prince George's County, 1970



Based on the 1970 Census of Population

Map 5.3 Mean Household Income in PG County, 1970



Based on the 1970 Census of Population tract level data

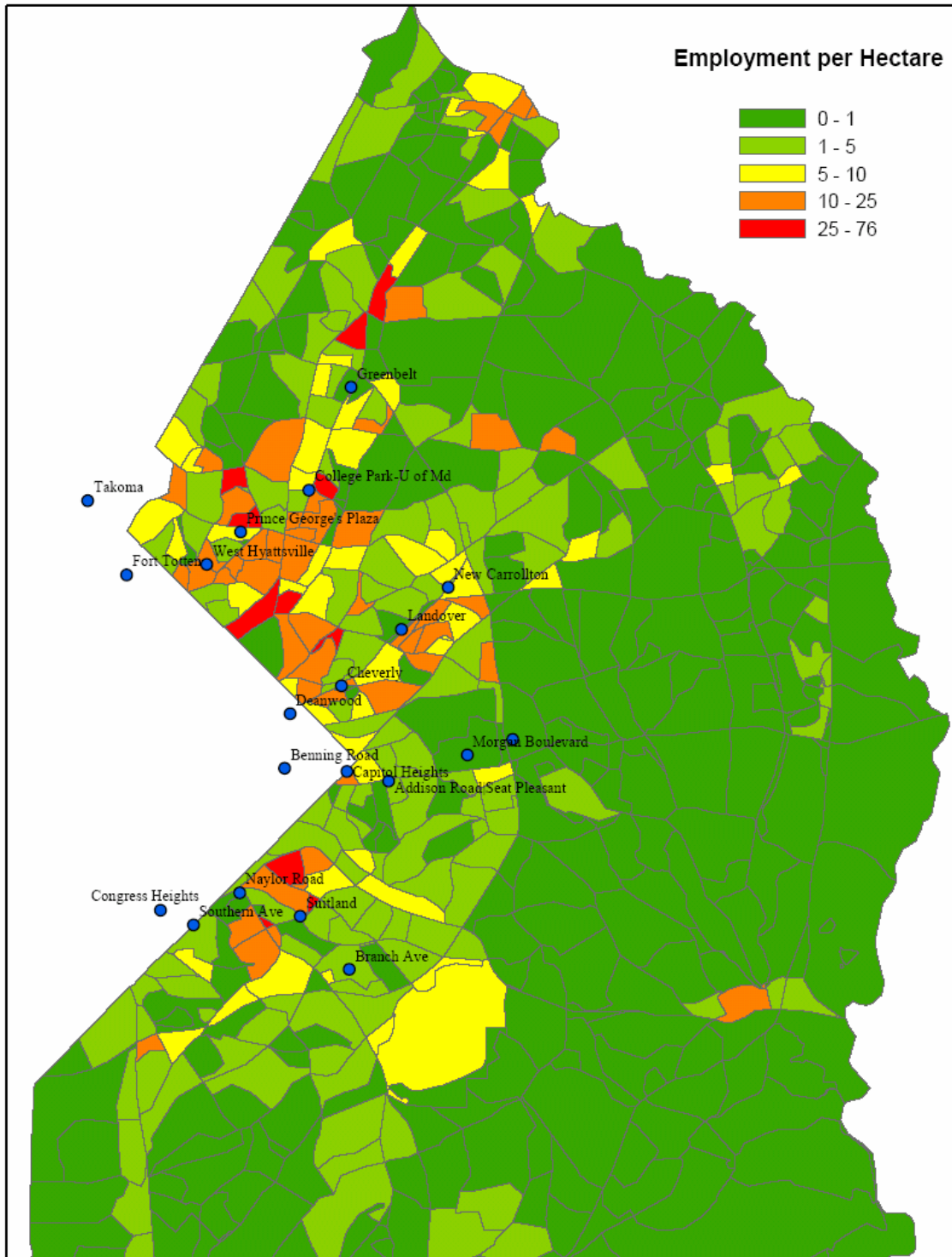
The employment density distribution was similar in Prince George's County as in Montgomery County (Map 5.4). As discussed in Chapter 3, the overall employment densities were lower in Prince George's. Again, most of the higher employment TAZs were near planned Metro stations and the District border. Some pockets of higher density employment also existed elsewhere; for example, in Upper Marlboro (east central part of the County), where the County seat is located, there was a TAZ with high employment density. The high population density areas were mainly low in employment density, suggesting the separation of land uses. Of the stations that opened before 2000, Greenbelt station was the only area with both low employment and population densities.⁷³ Morgan Boulevard and Largo Town Center stations were the most rural in nature, but both of these opened in the 2000s and are not part of the outcome analysis.

Comparing zoning in 1961 (Map 5.5) with actual land use in 1973 (Map 5.6) it is evident that in Prince George's County the zoned use in 1961 does not necessarily reflect actual land use in 1973. In comparison with Montgomery County, in Prince George's County more land was zoned for industrial and high-density residential use in 1961. This difference highlights the dissimilar type of the employment and population living in the County. As in Montgomery, most of the developed areas were around the District line. The farther lying parts of the County were for the most part zoned for rural uses. The land use around the proposed stations was mixed in 1973. Many stations had some area in commercial activity, but others were mainly in residential or rural uses. Based on the

⁷³ New Carrollton, Landover, Cheverly, Deanwood, West Hyattsville Prince George's Plaza, College Park, Greenbelt, Capital Heights and Addison Road

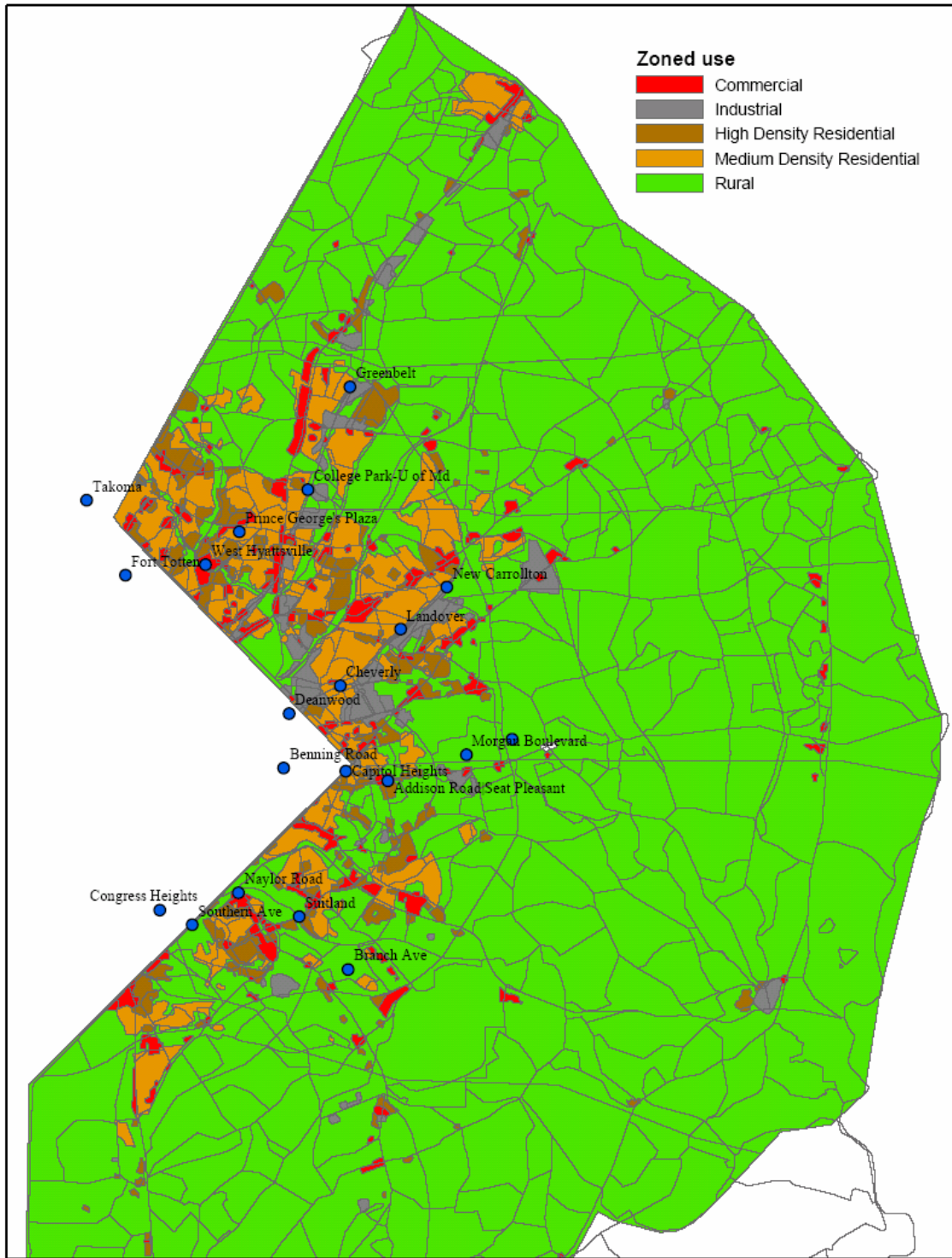
maps alone it is not possible to determine the impact of the various land uses on the decision of where to locate a station.

Map 5.4 Employment Density in PG County, 1972



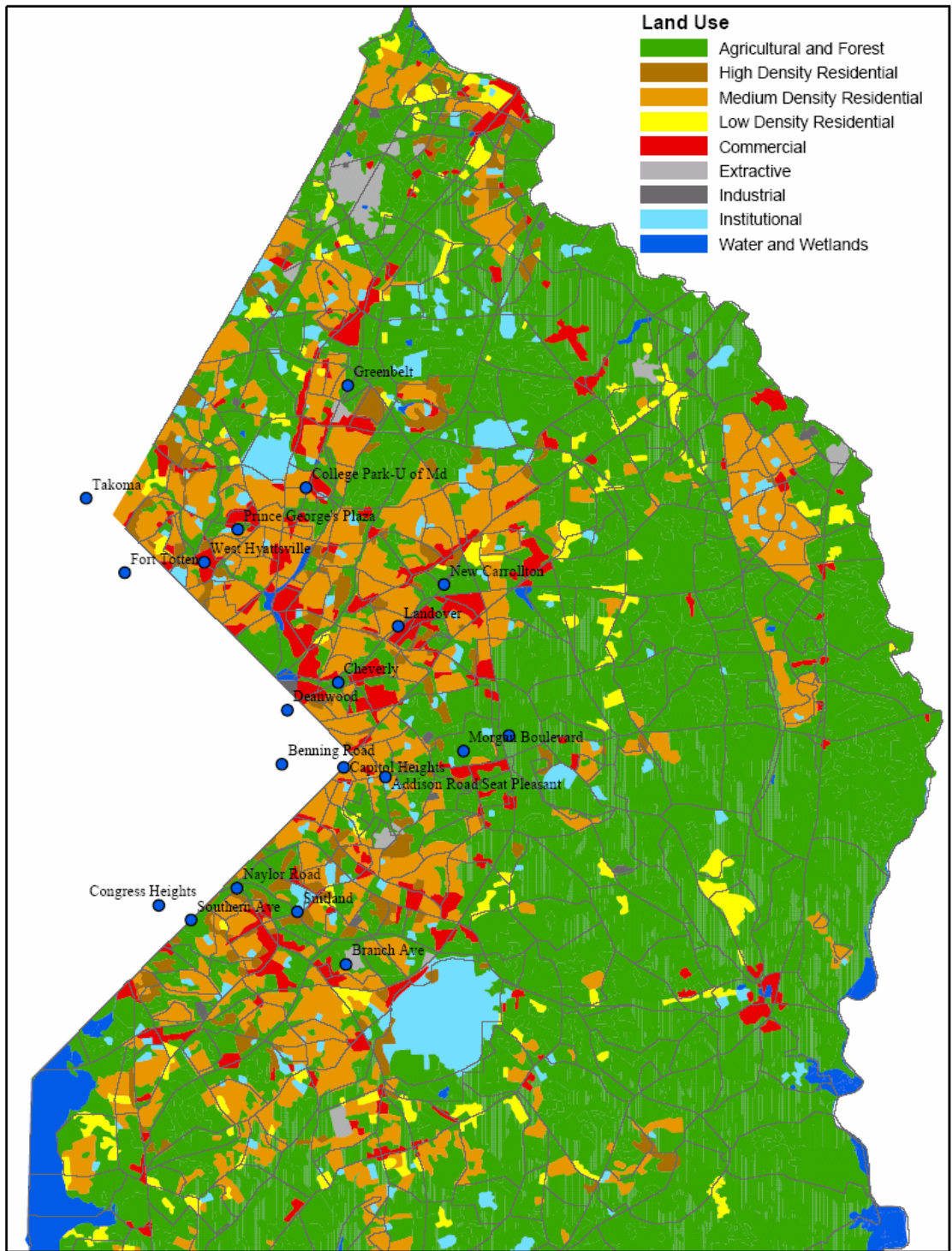
Based on 1972 aggregate TAZ level data from Green and James (1993)
Disaggregated TAZ level data are obtained by redistributing employment using land use information

Map 5.5 Zoning in PG County, 1961



Based on general zoning map in "Wedges and Corridors", MNCPPC, 1962

Map 5.6 Land Use in PG County, 1973



Based on Maryland Department of Planning land use map, 1973

The differences in the initial conditions are quantitatively captured in Table 5.1. The station areas are summarized by the year that they began operating. For all of the three sets of stations, the standard deviations are large for most of the variables suggesting a variety of different types of TAZs within the zone of influence of a Metro stop. In several aspects the three sets of TAZs are different. In general, the stations that opened after 2000 were more rural in character. The TAZs in station areas that opened earlier had less area zoned for low density residential uses in 1961.⁷⁴ Furthermore, the socio-economic composition was slightly different. The percent of white population was higher in the northern part of the Green Line (stations opened in 1993) than in the other two sets and lowest around the station areas that opened prior to 1990. These differences are statistically significant at the 99% level of significance. The pre-1990 stations were farther away, on average, from a major intersection than the TAZs surrounding stations built after 1990.⁷⁵

⁷⁴ The differences are statistically significant at the 99% level of significance.

⁷⁵ The differences are statistically significant at the 95% level of significance.

Table 5.1: Descriptive statistics of initial conditions around Prince George's Metro stations by year opened

	Station areas opened 1978 or 1980		Station areas opened 1993		Station areas opened in 2000s	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Pct of population white in 1970 ¹	0.56	0.35	0.94	0.06	0.81	0.15
Mean household income in 1969, (\$1000) ¹	9.36	4.85	10.01	6.13	10.34	4.62
Mean income (\$1000) squared in 1969	110.60	66.88	136.81	112.78	127.52	65.74
Population density in 1970 (people/ha) ²	23.41	21.15	28.91	30.66	16.50	21.00
Distance to closest intersection (km) ³	1.24	0.64	1.00	0.48	0.96	0.56
Distance to the White House (km) ⁴	12.53	2.05	12.65	3.07	12.65	2.98
Pct of area in agricultural use or forest (adjusted for protected lands) in 1973 ⁵	0.21	0.23	0.19	0.22	0.41	0.32
Pct of dwelling units built post 1960 in 1970 ⁶	0.53	0.41	0.64	0.39	0.49	0.40
Pct of dwelling units in apartments 1970 ⁶	0.20	0.30	0.23	0.33	0.31	0.38
Pct zoned residential high density 1961 ⁷	0.12	0.17	0.17	0.23	0.08	0.12
Pct zoned residential low density 1961 ⁷	0.25	0.31	0.27	0.30	0.65	0.34
Employment density in 1972 (workers/ha) ⁸	6.19	5.66	9.18	10.51	5.58	11.82
Observations	47		44		34	

Included in the station areas are those TAZs that have their centroid within a mile of a station. By 1990 the stations that were operational were Deanwood, Cheverly, Landover, New Carrollton, Capitol Heights and Addison Road. In the 1990s, four new stations opened - West Hyattsville, PG Plaza, College Park and Greenbelt. The stations that opened in the 2000s include Morgan Boulevard, Largo Town Center, Southern Avenue, Naylor Avenue, Suitland, and Branch Avenue.

¹ Based on tract level averages from the 1970 Census of Population

² Based on tract level averages from the 1970 Census of Population re-distributed using dwelling units information from the Maryland PropertyView database

³ Distance from the centroid of the tract to the closest major intersection as determined by 1971 Maryland highway maps

⁴ Distance from the centroid of the tract to the White House

⁵ Based on a Maryland Department of Planning land use map for 1973

⁶ Based on information extracted from the Maryland PropertyView 1999 and 2002 databases

⁷ Based on a MNCPPC (1961) zoning map

⁸ Based on Green and James (1993) employment data, redistributed to the TAZs using land use data

5.2 Propensity Scores

The variables summarized in Table 5.1 are used to determine the probability of a TAZ being within a mile of a station. That is, $T=1$ if the centroid of the TAZ is within a mile of a station. Using the standard normal distribution, the equation to be probability of receiving treatment is given by

$$P(T = 1|X) = \int_{-\infty}^{\beta'X} \phi(z) dz \quad (5.1)$$

where X includes the above-mentioned initial conditions thought to influence the location decision of Metro stations. Given the small number of TAZs within a mile of a station in 1990 (only 47 TAZs), I only calculate the propensity score (and the treatment impacts) for 2000.

The regression results for the propensity score are presented in Table 5.2. The model explains the siting decision well. The hypothesis that the coefficients are jointly equal to zero is rejected. (The $\chi^2(12)$ has a value of 131.)

Table 5.2: Propensity Score Calculation for Prince George's County

Variable	Year 2000	
	Coef.	Std. Err.
Pct of population white in 1970	-1.386	0.410 ***
Mean household income in 1969 (\$1000)	-0.122	0.061 **
Mean income (\$1000) squared in 1969	0.004	0.003
Population density in 1970 (people/ha)	-0.011	0.006 *
Distance to closest intersection (km)	-0.391	0.147 ***
Distance to the White House (km)	-0.142	0.024 ***
Pct of area in agricultural use or forest (adjusted for protected lands) in 1973	0.138	0.439
Pct of dwelling units built post 1960 in 1970	0.334	0.301
Pct of dwelling units in apartments in 1970	0.619	0.393
Pct zoned residential high density in 1961	-0.061	0.632
Pct zoned residential low density in 1961	-1.637	0.351 ***
Employment density in 1972 (workers/ha)	-0.040	0.013 ***
Constant	4.915	0.645
Number of observations	570	
coeff=0, Chi2 (12)	131 ***	

* significant at 10%; ** significant at 5%; *** significant at 1%

The propensity score is used to calculate the probability of a TAZ (centroid) of being within a mile of a Metro station.

There are some similarities in the siting decisions in Prince George's County and Montgomery County. As in Montgomery County, the closer the TAZ was to the White House or to a major intersection, the more likely it was to be within a mile of a station.

That is, accessibility from the TAZ was already better than average in the County. As expected, the larger the share of the area zoned for low density residential in 1961, the lower the probability of a Metro station eventually being built in the location. The above three results reflect the more urban nature of the TAZs where Metro stations were built.

There are also some differences in the siting decisions. The Metro stops in Prince George's County were built in areas with lower average household incomes (in contrast with the quadratic relationship found in Montgomery County). The negative relationship may come from the opposition in many of the more affluent neighborhoods to the siting of the Metro station in their jurisdiction. Alternatively, it is possible that the negative relationship reflects that Prince George's border with the District was predominantly low-income and any Metro line joining Prince George's with the District needed to go through the low-income areas. Furthermore, the percentage of population belonging to a minority is statistically significant in explaining Metro station location. The higher the share of the population that belonged to a minority, the higher the probability that a Metro stop was located in the area, holding all the other factors constant. This relationship suggests that there was no discernable discrimination against minorities in terms of station locations. Also, this finding is in contrast with the relationship found in Montgomery County where race played no role in the location decision.

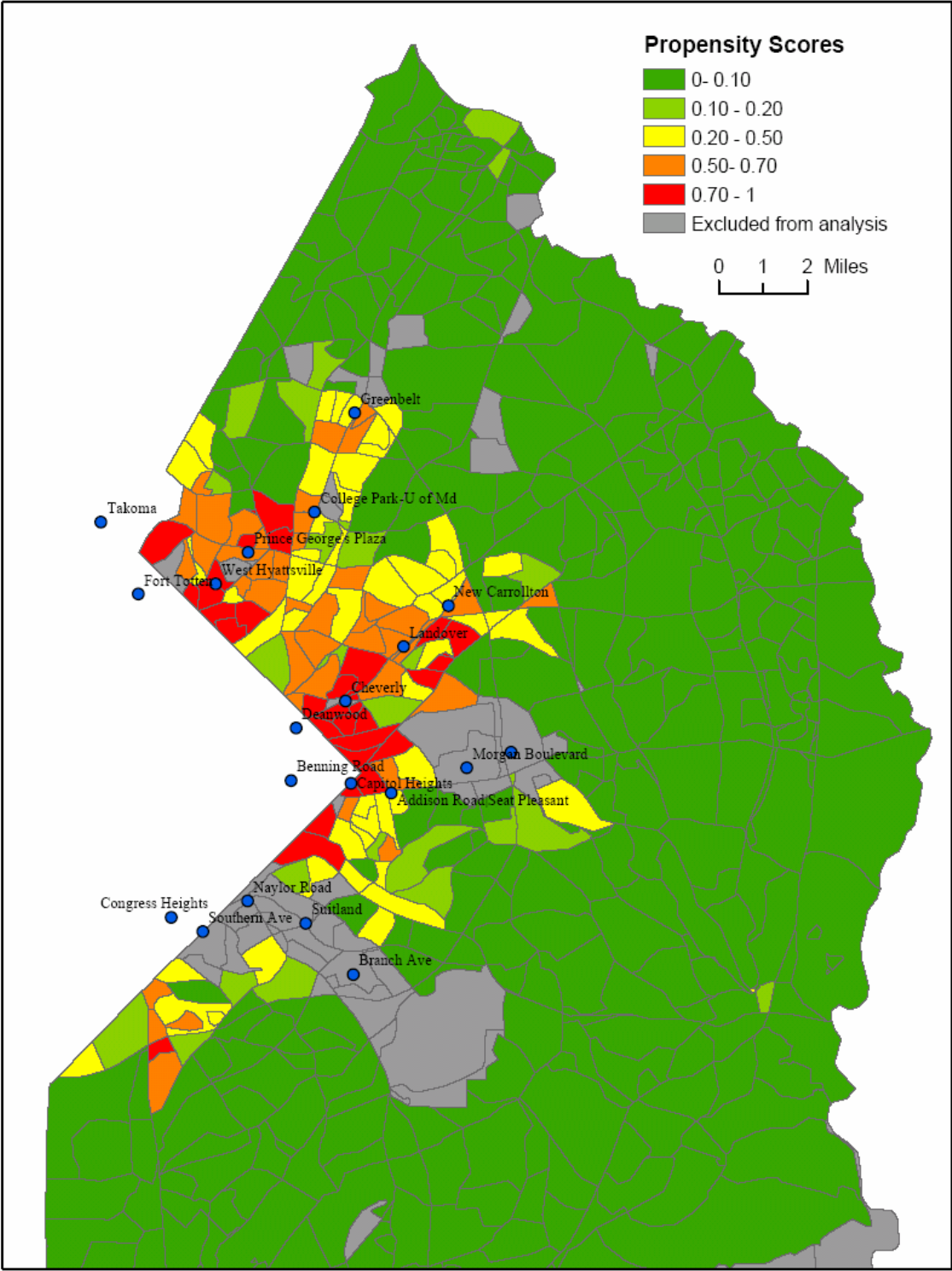
Curiously the station locations are *negatively* correlated with population density and employment density, *ceteris paribus*. These factors may reflect the general negative attitude towards the mass transport system in the County. Neighborhoods with higher

concentrations of population and employment and with the other favorable characteristics for a station location were able to persuade the officials not to locate the station within their vicinities. This relationship is contrary to what is observed in Montgomery County, where the Metro stations were placed in the high-density areas, *ceteris paribus*, and setting aside income restrictions, maximizing the potential number of users.

Map 5.7 spatially depicts the distribution of the propensity scores in Prince George's County. The higher propensity scores are all located around the District border. Also, in general, the treatment TAZs closer to the District boundary have higher propensity scores and those farther along the lines are more likely to have moderate propensity scores.

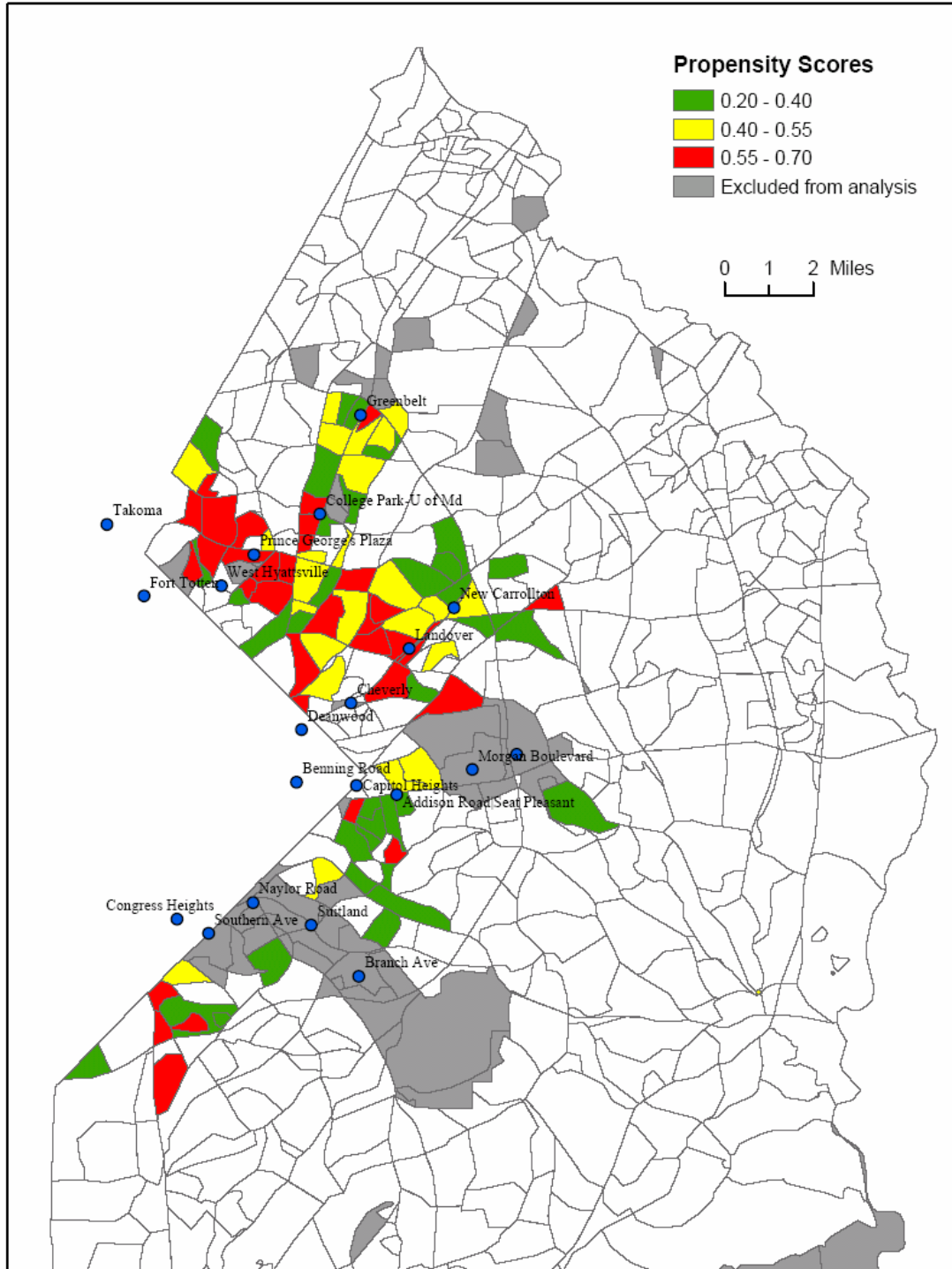
Table 5.3 and Figure 5.1 present the distribution of the propensity scores. Again, following Black and Smith (2004), I choose a range of propensity scores where there is a relatively high number of each type of observation—treatment and control. The propensity scores included in the thick support range from 0.2 to 0.7. Given the much smaller sample size for Prince George's County, the number of TAZs in each interval is small. With the above definition, the thick support sample has 47 treatment observations and 77 controls. In the case of Prince George's County most of the treatment observations excluded in the thick support sample are those adjacent to the District border (Map 5.8). Excluding from the thick support sample the observations with propensity scores between 0.2 and 0.3 or 0.6 and 0.7 does not change the treatment impacts significantly.

Map 5.7 Propensity Scores for PG County



Propensity scores calculated using the 2000 TAZ boundaries.

Map 5.8 Thick support TAZs in PG County



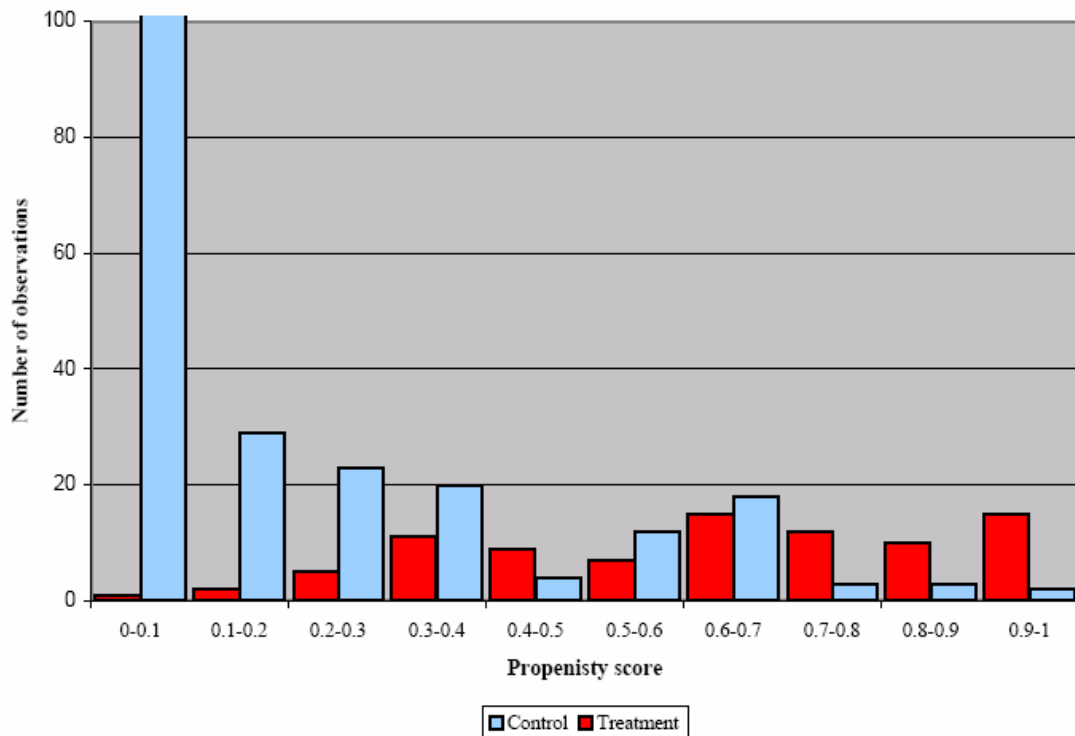
Propensity scores calculated using the 2000 TAZ boundaries.

Table 5.3: Propensity score distribution of TAZs in PG County, 2000 TAZ

Propensity Score range		Number of observations	
Max	Min	Control	Treatment
1	0.954		6
0.954	0.9	2	9
0.9	0.8	3	10
0.8	0.7	3	12
0.7	0.6	18	15
0.6	0.5	12	7
0.5	0.4	4	9
0.4	0.3	20	11
0.3	0.2	23	5
0.2	0.1	29	2
0.1	0	337	1

Highlighted rows make up the thick support region for the outcome analyses.

Figure 5.1: Distribution of propensity scores, Prince George's County, 2000



5.3 Construction of the counterfactual

The same estimation methodology used for Montgomery County is used to determine the treatment impacts in Prince George’s County (See section 4.3). That is, the matching estimator used is the Epanechnikov kernel with the optimal bandwidths determined by cross-validation (Black and Smith, 2004; Frölich, 2004a, 2004b). The Epanechnikov kernel is given by Equation 5.2.

$$W(i, j) = \frac{K\left(\frac{\hat{P}_j - \hat{P}_i}{h}\right)}{\sum_{k \in I_0} K\left(\frac{\hat{P}_k - \hat{P}_i}{h}\right)} \quad (5.2)$$

where h is the bandwidth of the kernel and \hat{P}_i and \hat{P}_j , are the probabilities of receiving treatment for a treatment observation i and a control observation j , respectively, I_0 is the set of possible control observations and $K(z) = 0.75(1 - 0.2z^2)/\sqrt{5}$ if $|z| < \sqrt{5}$ and is 0 otherwise. Table 5.4 gives the optimal cross-validated bandwidths based on the methodology outlined in Chapter 2. Eight different bandwidths based on the information from the non-station area TAZs are tested for each outcome measure.⁷⁶ The bandwidth with the lowest mean squared error is chosen as the optimal cross-validated bandwidth. As in Montgomery County, the optimal bandwidths range from 0.1 to 0.4.

⁷⁶ The bandwidths tested are 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, and 0.6.

Table 5.4: Optimal Bandwidth for Epanechnikov kernel for Prince George's County

Outcome measure	Outcome year	Cross-validated bandwidth
Employment density (workers/ha)	2000	0.2
Population density (people/ha)	2000	0.3
Development density (jobs+dwelling units/ha)	2000	0.2
Dwelling density using Census data (units/ha)	2000	0.3
Dwelling density using PropertyView data (units/ha)	2000	0.3
Dwelling density change 1970 to 2000 (units/ha)	2000	0.4
Dwelling density change 1990 to 2000 (units/ha)	2000	0.15
Percentage minority (%)	2000	0.1
Mean Household Income* (dollars)	2000	0.1

Optimal bandwidths are calculated using cross-validation, or leave-one-out, methods. The algorithm used for cross-validation was provided by Black and Smith which they used in Black and Smith (2004). The possible bandwidth tested were 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, and 0.6. The mean squared error did not vary greatly in this range.

Using on the weight matrix obtained from the Epanechnikov kernel with cross-validated bandwidths, it is possible to determine the “weighted” initial conditions in the control group. That is, each control observation is weighted by the total weight it has in the construction of the average counterfactual and this weight is applied to the vector of initial conditions. In order for the treatment and control group to be balanced in the vector of initial conditions, x , the weighted means of the two samples cannot be statistically significantly different. The optimal bandwidth is important in determining

whether or not the samples are balanced. When the bandwidth is larger, control TAZs with more dissimilar propensity scores are given relatively more weight (than when the bandwidth is smaller) and thus the differences in the initial conditions, in general, between the two groups also become larger. For these analyses the samples are balanced unless otherwise noted.

Table 5.5 summarizes the initial conditions in Prince George's County for both the common support sample as well as the thick support sample. The table reports both the unweighted as well as the weighted means of the variables. The weighted means are based on the Epanechnikov kernel with a bandwidth of 0.1. The weighted sample size for the treated observations is smaller since treatment observations i for which $\hat{P}_i > \max(\hat{P}_j) \forall j \in I_0$ are excluded from the sample. The tables also give the probability that the two sample means are equal.

As expected, when all the observations are used the *unweighted* means for the treated and non-treated are different. The Metro stations were located in areas that were more densely populated, closer to the White House and with a higher share of the land zoned for high-density development. In these cases the null hypothesis of equal means is rejected at 90 percent level of confidence. These differences suggest that any comparison of the groups that fails to account for these initial differences also fails to capture the true treatment impacts by not internalizing the endogeneity of the Metro location decision.

Table 5.5: Comparison of initial conditions in unweighted and weighted samples, PG 2000 TAZs

Variable	Sample	All observations			Thick support		
		Mean		p> t ¹	Mean		p> t ¹
		Treated	Control		Treated	Control	
Pct of population white in 1970 (%)	Unweighted	0.74	0.86	0.12	0.85	0.84	0.86
	Weighted	0.79	0.81	0.20	0.85	0.84	0.91
Mean household income in 1969 (\$1000)	Unweighted	9.96	12.48	0.13	10.89	10.37	0.67
	Weighted	10.17	10.04	0.88	10.89	10.69	0.85
Mean household income (in \$1000) squared	Unweighted	126.72	175.34	0.12	141.33	132.27	0.65
	Weighted	130.48	120.69	0.55	141.33	133.80	0.69
Population density in 1970 (people/ha)	Unweighted	27.27	9.19	0.07	30.66	23.87	0.39
	Weighted	27.79	27.78	0.84	30.66	27.26	0.61
Distance to closest intersection (km)	Unweighted	1.10	1.93	0.11	1.13	1.09	0.80
	Weighted	1.06	1.02	0.37	1.13	1.04	0.55
Distance to the White House (km)	Unweighted	12.44	20.91	0.05	13.32	13.15	0.80
	Weighted	12.57	11.74	0.21	13.32	12.56	0.35
Pct of area in agricultural use or forest (adjusted for protected lands) in 1973	Unweighted	0.19	0.60	0.06	0.23	0.34	0.26
	Weighted	0.20	0.22	0.41	0.23	0.29	0.39
Pct of dwelling units built post 1960 in 1970	Unweighted	0.60	0.46	0.19	0.63	0.53	0.38
	Weighted	0.60	0.58	0.59	0.63	0.55	0.45
Pct of dwelling units in apartments 1970	Unweighted	0.22	0.10	0.15	0.24	0.30	0.51
	Weighted	0.23	0.28	0.27	0.24	0.32	0.39
Pct zoned residential high density 1961	Unweighted	0.15	0.03	0.07	0.17	0.13	0.47
	Weighted	0.16	0.13	0.40	0.17	0.16	0.84
Pct zoned residential low density 1961	Unweighted	0.23	0.86	0.03	0.30	0.53	0.16
	Weighted	0.25	0.35	0.14	0.30	0.45	0.21
Employment density in 1972 (workers/ha)	Unweighted	7.35	2.60	0.11	7.97	6.67	0.55
	Weighted	7.48	6.75	0.51	7.97	7.19	0.69
Number of observations	Unweighted	87	451		47	77	
	Weighted	81	451		47	77	

Weights are based on an Epanechnikov kernel with a 0.1 bandwidth. The thick support propensity scores range from 0.2 to 0.7.

Calculated using *pstest* algorithm in Stata written by Edwin Leuven and Barbara Sianesi.

¹ Gives the probability of the equality of the treated and the controls using Hotelling's T-squared test for each propensity score quantile.

In the common support sample, weighting the control group observations using the Epanechnikov kernel weights yields similar initial conditions in the treatment and the control groups (Table 5.5). That is, by weighting we construct a counterfactual with statistically similar initial conditions to the treatment group.

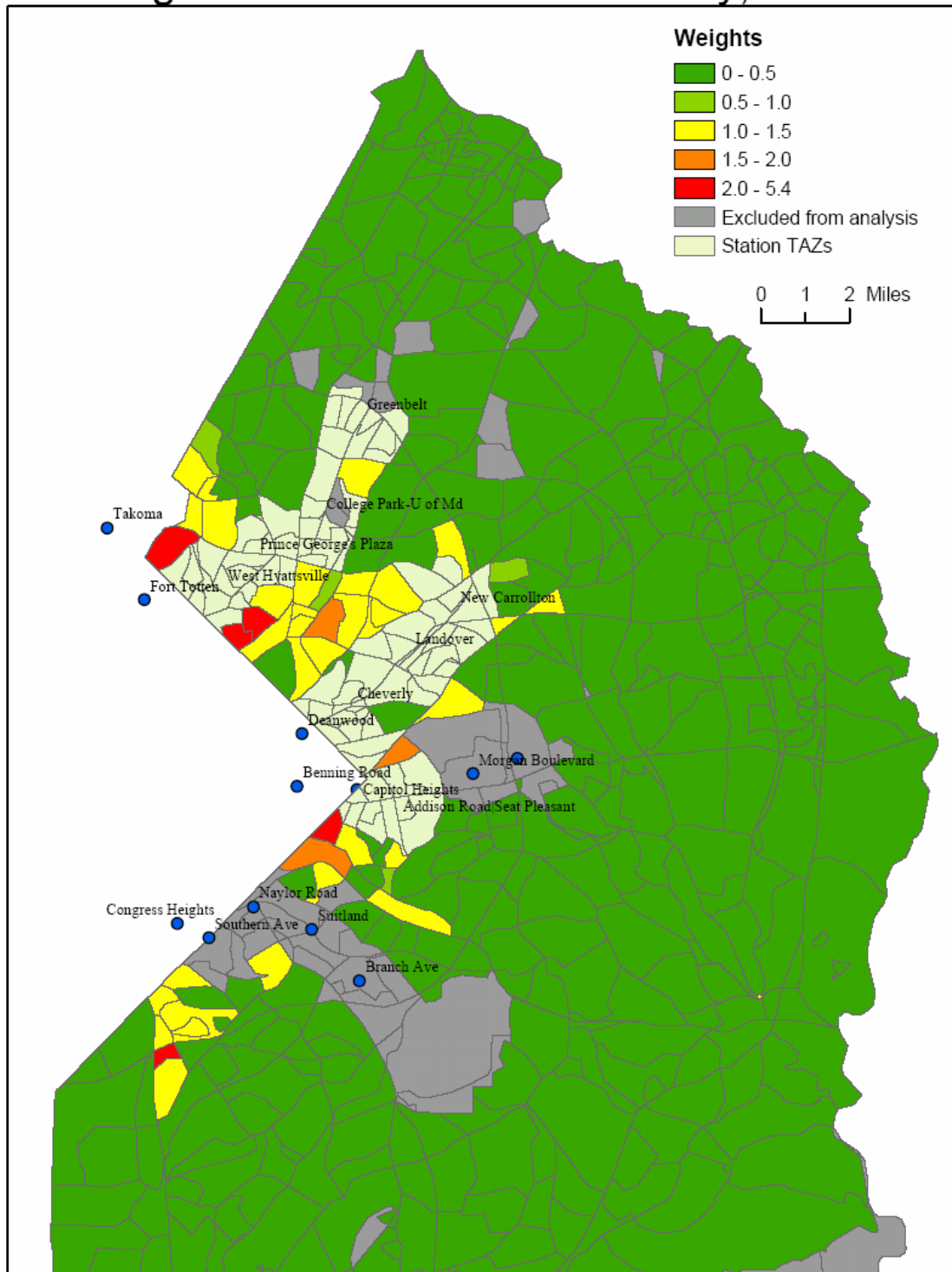
Comparing the initial conditions surrounding the station areas in Prince George's (Table 5.5) with those in Montgomery (Table 4.5) highlights some of the main differences. The population densities were similar, but the type of housing was quite

different in the two counties. Besides the more obvious difference in socio-economic characteristics of the two counties, in 1970 a slightly smaller share of dwelling units were built in the 1960s in Prince George's County than in Montgomery. That is, the housing stock was older in Prince George's. Furthermore, in Prince George's County the percentage of total dwelling units in apartments was greater than in Montgomery County reflecting the lower income in the Prince George's County. Moreover, a larger percentage of land area was in agriculture or as forest in 1973. Prince George's County had pockets of high-density housing surrounded by still undeveloped land. Montgomery County had much more of its station areas already devoted to developed uses by 1973.

Map 5.9 shows an example for the common support sample of the weights used to construct the counterfactual. The map is based on weights obtained when the Epanechnikov kernel bandwidth is set to 0.1. Again most of the weight in the counterfactual is given to TAZs that are adjacent to the treatment TAZs. It is interesting to note, that the highest weight given to any control TAZ is less than 5.5, whereas in Montgomery there were control TAZs that had weights of nine. These differences in the weights indicate that the counterfactual for Prince George's County does not rely as heavily on a few control TAZs and that in some sense the differences between the treatment and control groups are not as marked.

Map 5.9

Weights for controls in PG County, 2000



The weights are based on an Epanechnikov kernel with a bandwidth of 0.1

Similar to Montgomery County, once only the observations in the range of thick support are used, the *unmatched* initial conditions are statistically similar between the control and treatment groups. With the matching these differences are, in general, reduced further (Table 5.5). Again, this suggests the superiority of the thick support sample in terms of comparing similar samples with the caveat of a reduced number of observations and a loss of generality of countywide impacts. In the case of Prince George’s County the reduced number of treatment observations becomes of greater concern than in the Montgomery case.⁷⁷

5.4 Impacts of Metro stations in Prince George’s County

The treatment impacts are obtained using Equation (5.3) and discussed in detail in Chapter 2. The average treatment impact can be written as,

$$\hat{M}(T) = \sum_{i \in I_1 \subset CS} \omega(i) [y_i^1 - \sum_{j \in I_0} W(i, j) y_j^0] \quad (5.3)$$

where $\omega(i) = 1/N_1$ and N_1 is the number of treated observations and the W matrix is given by Equation (5.2). The outcome variables of interest are: employment, population and overall development densities as well as the income and racial makeup of the TAZs. Following I give the treatment impacts for 2000 for both the one-mile radius as well as the half a mile radius. Also, I calculate the impacts for both the whole sample as well as the “thick support” sample described above.

In Prince George’s County, there are positive treatment impacts on employment and overall density in 2000 in the various samples and treatment distances used. There is

⁷⁷ However, in both cases we are left with slightly more than half of all of the treatment TAZs and about 15 percent of the total possible control TAZs.

also some evidence of a negative impact on the minority population, although there is no evidence of a negative impact on average household income.

The first set of columns in Table 5.6 present the differences in the various outcome measures between all the treatment TAZs and the unweighted controls. That is, comparing the mean of those TAZs within a mile of a station to the mean of those TAZs that are farther than a mile from a station. Under the assumption that all TAZs are equally likely to be within a mile of Metro stop, that is, siting is exogenous to the initial conditions, we observe great differences in the outcome measures in 2000. The means of the two samples are statistically different except for two measures—change in the number of dwelling units from 1970 to 2000 and the percentage of the population belonging to a minority. Comparing with the results from Montgomery County (Table 4.7), the 2000 average densities in the TAZs within a mile of a station are much lower and concurrently the unmatched impacts are also lower in Prince George’s County. The average employment density in a TAZ within a mile in Montgomery is close to 42 employees per hectare. In Prince George’s it is only around 11 workers per hectare. The employment density for the non-station area TAZs is approximately the same in Prince George’s and Montgomery. Similarly, the population density in Montgomery was approximately twice that of Prince George’s County in the TAZs within a mile of a station. For the rest of the TAZ the population density was approximately the same in the two counties.

Table 5.6: Average treatment impacts for one mile radius in 2000 for Prince George's County

	No matching ¹			All observations within Common Support				
	Average treated	Average control	Unmatched impact	Average treated	Average matched control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound
Employment density (workers/ha)	11.24	3.64	7.59	11.35	5.98	5.37	2.18	8.89
Population density (people/ha)	20.81	12.74	8.07	22.61	24.83	-2.22	-7.99	3.58
Development density (jobs+dwelling units/ha)	17.40	6.81	10.60	18.64	14.28	4.36	0.73	7.83
Dwelling density using Census data (units/ha)	8.29	4.78	3.52	8.98	9.25	-0.27	-2.67	2.17
Dwelling density using PropertyView data (units/ha)	7.81	3.89	3.92	8.44	9.01	-0.57	-2.55	1.88
Dwelling density change 1970 to 2000 (units/ha)	0.95	1.69	-0.74	1.04	1.59	-0.55	-0.92	-0.11
Dwelling density change 1990 to 2000 (units/ha)	0.27	0.44	-0.17	0.29	0.43	-0.14	-0.30	0.10
Percentage minority (%)	0.59	0.53	0.06	0.61	0.74	-0.13	-0.23	-0.01
Mean Household Income (dollars)	40,989	65,861	-24,872	42,990	47,088	-4,098	-11,179	2,529
Number of observations	91	498		78	451			

The treated observations are those TAZs that have their centroid within a mile of a station. All other TAZs are controls. The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 5.4 and using *psmatch2* algorithm in Stata. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval.

Statistically significant impacts are in boldface.

¹ All of the "No matching" impacts are statistically significant at the 95% level of confidence, except for the dwelling density change from 1990 to 2000 and the percentage of population belonging to a minority.

Using the Epanechnikov kernel to weight the control observations, there is some evidence of a positive treatment impact on employment density and on overall development density in 2000 within the one-mile radius (Table 5.6). The impact on employment density is approximately 5 additional workers per hectare. The impact on overall development density is of the same magnitude. These impacts are less than one-third of the impact in Montgomery County. Unlike in Montgomery County, in Prince George's County the point estimate for the overall development density impact is actually slightly smaller than the point estimate for the impact on employment. This is explained by the fact that dwelling units are not more likely to be located in the treatment TAZs.

Even though the unweighted differences indicate dissimilarities in the mean population and dwelling unit densities there is no evidence of positive impacts on population and dwelling unit densities from the Metro stations after weighting the control observations. In fact, there were slightly more (0.5 units per hectare) new dwelling units built between 1970 and 2000 in the control TAZs than in the treatment TAZs. This negative impact may be due to the fact that Prince George's planners were encouraging development of areas such as Laurel and Bowie, well beyond the reach of the Metro system (Green and James, 1993). The negative impact could also be from perceived disamenities from the Metro stations. Furthermore, the overall growth in population over the three decades in Prince George's County was only approximately 21 percent. The pressure for new development was much lower than in Montgomery County, which grew around 50% in population the same span of time.

Another difference between Montgomery and Prince George's counties is that in Prince George's County minorities appear to have been negatively affected by the Metro system. On average, in the control TAZs minorities make up 74% of the population whereas in the TAZs within a mile of a station they only make up 61% of the population. The difference is statistically significant. The average share of minorities in 1970 for the treated and control TAZs were 21% and 19%, respectively. There was a large immigration of minorities into Prince George's County since the opening of the system. However, the minority population was more likely to find housing in areas beyond the one-mile radius of a Metro station. However, in Prince George's County there was no impact on the average household income. The Metro areas do not attract higher income households.

When only the thick sample TAZs are considered, the results remain fairly similar to those in the common support sample (Table 5.7).⁷⁸ There is an impact of about 7 employment opportunities from being within a mile of a subway station. This translates into about 5,690 additional jobs around a subway station or 51,213 additional jobs around the 9 Metro stations in operation in 2000. These jobs maybe jobs that relocated due to the Metro or new jobs that came into the County because of the Metro. This is about 17% of the total jobs in Prince George's County. Again, there is no impact on population density, dwelling density or average household incomes. In the thick support region we continue to observe the negative impact on minorities from the Metro system. The control TAZs consists of around 76% minority population whereas the treatment areas

⁷⁸ If the thick sample is confined to those propensity scores between 0.3 and 0.7 (excluding propensity scores between 0.2 and 0.3 from the sample), the results do not change significantly.

only have a minority population of about 59%. So even when only those TAZs that had a moderate chance of receiving a station are considered the negative impact on minorities persists.

Unlike in Montgomery County, we do not see great differences in the impacts when the sample is limited to the subset of TAZs within half a mile of a station (Table 5.7). The impacts on employment and total development densities become slightly larger but the differences in the impacts are not statistically significant. The impact on total employment around a station area is 2,164 jobs that would have located elsewhere than within half a mile of a station or 19,476 jobs around the 9 Metro stations. This represents about 7% of the total employment in the County in 2000. Even in the half a mile sample there are no statistically positive impacts on population density or the total dwelling unit density. The change in the dwelling unit density both between 1970 and 2000 and between 1990 and 2000 was slower in the treatment TAZs than in the control TAZs. These results could arise if construction for office space around the stations is displacing housing construction or if there are negative amenities from the Metro stations which deter residential demand for the locations. The negative impact on minorities increases slightly to a difference of 22 percentage points, but again is not statistically significantly different from the impact in the one-mile sample.

Table 5.7: Average treatment impacts for thick support TAZs in 2000 for Prince George's County

	Thick support, treatment one mile radius				Thick support, treatment half a mile					
	Average treated	Average matched control	Treatment impact	90% C.I. Lower Bound	90% C.I. Upper Bound	Average treated	Average matched control	Treatment impact	90% C.I. Lower Bound	90% C.I. Upper Bound
Employment density (workers/ha)	14.22	7.12	7.10	1.91	12.73	17.77	7.12	10.65	0.81	23.58
Population density (people/ha)	23.73	23.91	-0.18	-7.48	8.44	29.22	22.96	6.26	-7.20	28.67
Development density (jobs+dwelling units/ha)	21.48	14.76	6.72	1.77	12.58	25.90	14.34	11.56	2.04	24.18
Dwelling density using Census data (units/ha)	9.52	9.08	0.44	-2.78	4.14	12.14	8.75	3.39	-2.59	12.89
Dwelling density using PropertyView data (units/ha)	9.31	9.23	0.08	-2.90	3.67	11.40	8.95	2.45	-3.06	11.44
Dwelling density change 1970 to 2000 (units/ha)	1.09	1.72	-0.63	-1.20	0.01	0.93	1.74	-0.81	-1.43	-0.01
Dwelling density change 1990 to 2000 (units/ha)	0.27	0.50	-0.23	-0.48	0.08	0.17	0.49	-0.32	-0.58	-0.03
Percentage minority (%)	0.59	0.76	-0.17	-0.27	-0.05	0.53	0.75	-0.22	-0.40	-0.03
Mean Household Income (dollars)	44,735	48,334	-3,599	-11,058	4,227	42,971	48,495	-5,524	-19,031	6,476
Number of observations	47	77				15	77			

The treated observations are those TAZs that have their centroid within a mile of a station. All other TAZs are controls. The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 5.4 and using *psmatch2* algorithm in Stata. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval. Statistically significant impacts are in boldface.

The average treatment impacts in the thick samples are relatively robust to the definition of the thick sample. For the one mile radius when the sample includes those TAZs with propensity scores between 0.2 and 0.6 or those TAZs with propensity scores between 0.3 and 0.7 the results are qualitatively similar. For the half-mile sample with the changes in the definition of the thick sample, the negative impact on minorities ceases to be statistically significant, although the point estimate is of similar magnitude. The impacts on employment and development density as well as the changes in dwelling densities remain qualitatively similar.

In Prince George's County thick support samples there does not appear to be spillover impacts. Table 5.8 presents the treatment impacts when those control TAZs that are between one mile and 1.5 miles from a station are excluded from the analyses. The point estimates and the statistically significant treatment impacts remain qualitatively similar to those obtained when no buffer is used.

Table 5.8: Average treatment impacts for thick support TAZs in 2000 for Prince George's County (possible control TAZs are at least 1.5 miles from the closest station)

	Thick support, treatment one mile radius				Thick support, treatment half a mile					
	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound	Average treated	Average weighted control	Treatment impact	90% C. I. Lower Bound	90% C. I. Upper Bound
Employment density (workers/ha)	14.39	7.36	7.03	1.44	13.53	17.77	7.06	10.71	0.57	24.68
Population density (people/ha)	23.93	26.71	-2.78	-11.77	6.13	29.22	25.60	3.62	-10.26	26.26
Development density (jobs+dwelling units/ha)	21.74	16.25	5.49	0.03	11.46	25.90	15.42	10.48	0.38	24.65
Dwelling density using Census data (units/ha)	9.60	10.19	-0.59	-3.94	3.63	12.14	9.78	2.36	-3.89	12.45
Dwelling density using PropertyView data (units/ha)	9.45	10.24	-0.79	-4.41	2.85	11.39	9.87	1.52	-4.02	10.61
Dwelling density change 1970 to 2000 (units/ha)	1.12	1.82	-0.70	-1.30	-0.05	0.93	1.81	-0.88	-1.54	-0.12
Dwelling density change 1990 to 2000 (units/ha)	0.28	0.56	-0.28	-0.55	0.08	0.17	0.53	-0.36	-0.68	-0.06
Percentage minority (%)	0.58	0.77	-0.19	-0.31	-0.07	0.53	0.77	-0.24	-0.42	-0.05
Mean Household Income (dollars)	44,805	46,652	-1,847	-9,846	5,841	42,971	4,747	38,224	-17,950	7,713
Number of observations	46	56				15	56			

The treatment impacts are determined using an Epanechnikov kernel with bandwidths given in Table 4.4 and using *psmatch2* algorithm in Stata. Confidence intervals based on 2000 bootstrap repetitions, using the bias-adjusted interval. Statistically significant impacts are in boldface.

5.5 Conclusions

In general, in Prince George's County the Metro stations have positively impacted employment density. This impact also drives the significant impact in the overall development density measure. This densification result is consistent throughout the various samples used. There is no observed impact on the population density or dwelling unit density measures. Nor are differential impacts observed in the distribution of population and employment depending on the distance from the closest Metro stop (within half a mile versus within a mile). There is some evidence of minorities locating disproportionately in the non-station areas.

The treatment impacts in Prince George's County are, in general, smaller than those observed in Montgomery County. One explanation for these differences arises from the fact that the initial countywide conditions as well as the policies adopted after the Metro system opened were different. Also, the growth rates in the two counties were different (21% in Prince George's County versus 50% in Montgomery County between 1970 and 2000) potentially explaining differences in the observed impacts. In Prince George's County the pressure for new housing was much lower than in Montgomery County. From both counties there is some evidence of a negative impact on minorities. The impact is clearer in Prince George's County than in Montgomery. However, in Montgomery County there is also evidence of higher average household incomes near a Metro station. In Prince George's County such impact is not observed. Again, the impacts are qualitatively similar to the choice of the definition for the thick support sample and no substantial spillover impacts are found.

Chapter 6

Conclusions

Mass transportation networks—such as subways—are thought to impact the location decisions of households and employers, such that station areas become denser in terms of both population and employment. However, the empirical evidence of such densification is mixed. In this dissertation I analyze the impacts of the Washington Metro rail system on development patterns in Montgomery and Prince George’s Counties in Maryland using propensity score matching techniques. This methodology allows me to explicitly model the endogeneity in the station siting decision as well as relax functional form assumptions on the relationship between density and the factors thought to determine density. These techniques have not previously been applied in the impact analysis of transportation networks.

I calculate the impact of the proximity to a Metro station on a variety of measures including employment, population and dwelling unit density as well as the percentage of population belonging to a minority and average household income. I use transportation analysis zones (TAZs) as the unit of analysis and consider the treatment group to be composed of those TAZs that are within a mile of a station. Given that the Metro stations were not randomly located in the metropolitan area, it necessary to first model the station location decision. To this end, I calculate the probability of TAZs being with a mile of station based on the characteristics—such as, employment density, population density, land use and zoning—of the TAZs before the Metro was built.

Given the differences in the probability distributions of the TAZs in the treatment and in the control groups, I not only calculate the treatment impacts using the whole common support sample but I also define a thick support sample. The thick support sample consists of those observations that have a propensity score between 0.2 and 0.7—a range where “there are a substantial number of observations in both the treatment and comparison groups” (Black and Smith, 2004, pg. 118) such that no one control observation is heavily relied upon in the construction of the counterfactual. Furthermore, the use of the thick support sample is warranted, even though the two groups in the common support sample are statistically similar in the pre-Metro characteristics after the samples are weighted appropriately, since for some characteristics the statistical equality is narrowly rejected. Furthermore, to determine temporal and spatial facets of the impacts, I calculate the impacts for both 1990 and 2000 as well as for TAZs within a half a mile of a station. It must be stressed that the methodology cannot attribute the observed impacts to be caused by the Metro station alone but the impacts also incorporate any changes in decisions due to changes in land use planning measures in these areas.

The results of the analysis in this dissertation suggest that the two counties differed in their response to the construction of the Metro system and concurrent changes in land use policies. In addition, the gains from using propensity score matching methods instead of OLS differ.

In Montgomery County, the estimated impacts in the common support sample are greater than those in the thick support sample. This suggests there may be heterogeneous treatment impacts. For example, it is possible that the treatment impacts are higher for higher propensity scores and thus the treatment impact in the common support sample is higher than when the tails of the propensity score distribution are excluded from the analysis. It is also possible that there are unobservables that determine selection and that these have a greater influence on the tails of the distribution (Black and Smith, 2004). The thick support region results are less influenced by these factors. On average, even without weighting, the treatment and control groups are more similar in their initial conditions in the thick support sample. Therefore, I concentrate on the thick support region results. For Prince George's County the impacts based on the two samples do not differ significantly, suggesting that the tails of the probability distributions do not influence the results to the same degree as in Montgomery County. However, I concentrate on the thick support sample given that the two samples are more balanced in the smaller sample and this also allows me to compare the treatment impacts in the two counties.

The most consistent treatment impact is on employment density and on overall development density, but even these impacts are statistically insignificant for some samples.⁷⁹ When the thick support sample TAZs within a mile of a station are considered in the treatment group, the impact on employment is much greater in Prince George's County than in Montgomery County (7 additional workers per hectare versus no

⁷⁹ Overall density is the sum of dwelling unit density and employment density.

significant impact).⁸⁰ However, when only those stations within half mile of a station are considered, both treatment impacts are statistically significant (11 additional workers per hectare in Prince George's County and 15 additional workers per hectare in Montgomery County) suggesting higher impacts closer to the stations. Contrary to the one-mile impact, the half-mile impact is smaller in Prince George's County than in Montgomery County. These impacts translate into 7% and 9% of the total employment in the counties located within a half a mile of a station, respectively.

The impacts on overall development density reflect the impacts on employment density. That is, in the thick support sample the impacts are greater in the half mile treatment group than in the one mile treatment group. Also the impact in the one mile treatment group is greater in Prince George's County whereas it is greater in Montgomery County for the half mile treatment group.

The results for Montgomery County for different outcome years suggest that the impacts on employment and overall development are not immediate and long periods of time are necessary to observe changes in densities even if prices react before stations are in operation. In the sample of stations that opened in 1984 there was no evidence of a positive impact on employment density or overall development density in 1990, whereas in 2000 such impacts were statistically significant.

⁸⁰ There is a positive treatment impact on employment density in both Montgomery County and Prince George's County when the common support sample is used.

The impacts on dwelling unit density and changes in the dwelling unit density are distinct in the two counties studied. I find in Montgomery County a positive impact on dwelling unit density from proximity to a Metro station. This is also reflected in the fact that more of the new housing units were built in the station areas than in comparable areas farther away. In Prince George's County such densification impact on dwelling units is not observed and, moreover, the new housing units tend to be built in the non-station areas.

In terms of population there is no evidence of a positive impact on population density from the Metro system. However, there is a negative impact on percentage of minorities in the treatment areas in Prince George's County. This negative result is especially striking given that the share of minorities in the county's population has increased drastically from 1970 to 2000.

Besides the differences in the treatment impacts, the two counties differ in the gains that are realized when propensity score matching techniques are used versus OLS. First, any differences in the estimates based on matching using the common support sample and based on OLS using all the observations are due to the linear functional form assumption and the endogeneity of station location. The OLS sample may be larger than the common support sample. In general, no additional information is used in the OLS to limit the sample to a range of observations where both station and non-station TAZs exist as is done in matching estimators with the common support condition imposed. Second, the treatment impacts from the two methodologies should differ mainly due to the linear

functional form assumption in the thick support sample. In the thick support sample two samples are balanced without weighting, and the endogeneity of the station location has been accounted for.

Table 6.1 summarizes the treatment impacts for both, when an Epanechnikov kernel matching estimator with cross-validated bandwidths is used and when an OLS regression is run using both the common support sample and its OLS counterpart as well as the thick support sample. It is evident from these results that the OLS estimates are very different from the treatment impacts when a matching estimator is used. For Montgomery County, in the common support sample, the OLS estimates tend to be smaller than the estimates from matching for those treatment impacts that are statistically significant. That is, the OLS estimates underestimate the true treatment impact. On the other hand, the OLS analysis yields statistically significant treatment impacts for several outcome measures—such as percent minority and average household income—where no statistically significant impact is found once the endogeneity of station location is acknowledged and the functional form specification relaxed. For the thick support sample the results are more similar, suggesting that most of the differences in the common support sample are from endogenous station location and not from the functional form restriction.

For Prince George's County the results are slightly different. The impacts using OLS or using kernel matching are in general very similar, both in terms of the point estimates as well as statistical significance, regardless of the sample used. There are

consistently statistically significant impacts on employment density, development density and percent minority. It is not evident why endogeneity and functional form assumptions do not significantly alter the estimated treatment impacts in the Prince George's County analysis.

The analysis in this dissertation provides some evidence of densification from mass transit infrastructure networks especially in terms of employment and overall development in the long run. The research also suggests the potential problems from endogeneity of station location and a linear functional form assumption on the impact estimates. Additional mass transit systems need to be analyzed to determine the universality of these results. Furthermore, cross county analyses are necessary to determine to what degree complementary land use planning measures are affecting the results.

Table 6.1a: Comparison of impacts from OLS and propensity score matching, Montgomery County, 2000

	Employment density	Population density	Development density	Dwelling density based on Census	Dwelling density based on PropertyView	Change in dwelling density between 1970 and 2000	Change in dwelling density between 1990 and 2000	Percent minority	Mean household income
Epanechnikov kernel common support sample	16.87 (5.59)	13.63 (9.25)	21.98 (7.74)	5.28 (2.72)	3.18 (1.39)	3.41 (1.25)	0.83 (0.34)	-0.04 (0.03)	6,703 (4,588)
OLS whole sample	6.23 (3.36)	8.58 (5.92)	8.89 (4.16)	2.74 (1.81)	0.97 (1.10)	0.87 (1.11)	0.48 (0.27)*	-0.06 (0.02)	13,175 (5,520)
Epanechnikov kernel thick support sample	1.51 (3.22)	7.67 (7.79)	3.11 (3.57)	1.71 (1.73)	0.91 (1.74)	1.89 (1.47)	0.38 (0.27)	-0.02 (0.03)	13,790 (7,387)
OLS thick support sample	2.22 (2.50)	4.72 (7.14)	4.50 (2.80)	2.11 (1.35)	2.50 (1.16)	2.45 (1.17)	0.48 (0.26)	-0.06 (0.03)	13,174 (5,462)

Standard errors given in parenthesis. For the estimates based on an Epanechnikov kernel with bandwidths given in Table 4.4 the standard error is obtained by bootstrap methods with 2000 replications. The standard error reported here is the normal approximation.

Table 6.1b: Comparison of impacts from OLS and propensity score matching, Prince George's County, 2000

	Employment density	Population density	Development density	Dwelling density based on Census	Dwelling density based on PropertyView	Change in dwelling density between 1970 and 2000	Change in dwelling density between 1990 and 2000	Percent minority	Mean household income
Epanechnikov kernel common support sample	5.37 (2.10)	-2.22 (3.65)	4.36 (2.21)	-0.27 (1.48)	-0.57 (1.33)	-0.55 (0.25)	-0.14 (0.11)	-0.13 (0.06)	-4,098 (4,238)
OLS whole sample	4.06 (1.380)	-2.63 (1.597)	2.50 (1.43)	-0.60 (0.71)	-0.97 (0.56)	-0.64 (0.48)	-0.15 (0.15)	-0.18 (0.04)	-7,116 (4,457)
Epanechnikov kernel thick support sample	7.10 (3.34)	-0.18 (4.75)	6.72 (3.32)	0.44 (2.11)	0.08 (1.99)	-0.63 (0.37)	-0.23 (0.17)	-0.17 (0.07)	-3,599 (4,627)
OLS thick support sample	8.10 (2.65)	-1.26 (2.27)	6.98 (2.60)	0.12 (0.97)	-0.07 (0.76)	-0.40 (0.39)	-0.06 (0.12)	-0.21 (0.05)	-6,356 (2,605)

Standard errors given in parenthesis. For the estimates based on an Epanechnikov kernel with bandwidths given in Table 4.4 the standard error is obtained by bootstrap methods with 2000 replications. The standard error reported here is the normal approximation.

Appendices

Appendix 3.1: Tax exempted properties in Maryland PropertyView files included in the dwelling unit calculations

Exemption code	Description	Kept for residential dwelling unit calculations
100	Office Buildings	
110	Hospitals	
120	Parks	
130	Military Installations	
140	Schools	
150	Non-Military Airports	
160	Research Institutions	
170	Other	
180	Seized Properties	
190	Game Preserves	
200	Office Buildings	
210	Hospitals and Health Related Facilities	
220	Parks	
230	Police Stations and Barracks	
240	Armories	
250	Colleges	
260	Airports (Baltimore-Washington International)	
270	Museums	
280	Detention Centers	
290	Game Preserves	
300	Port Authority	
310	Other	
320	Other	
330	Department of Public Works	
335	North East Maryland Waste Disposal Authority	
340	Market Authority	
350	Other	
360	Metropolitan Transit Authority	
370	Housing and Urban Development	yes
	State Roads Commission (Mass Transit Administration)	
380	Administration)	
390	Tobacco Warehouses	
400	Office Buildings	
410	Hospitals and Health Related Facilities	
420	Parks and Recreation	
430	Police Stations	
440	Public Schools including Junior Colleges	
450	Airports	
460	Museums	
470	Detention Centers	
480	Off-Street Parking	

490	Fire Departments	
500	Public Works Properties	
501	Flood Plains	
502	Storm Drains	
503	Common Areas by Plat	
504	Open Spaces by Plat	
505	Flood Plains by Plat	
506	Landfills	
507	Wastewater Pumping Stations	
508	Freshwater Pumping Stations	
510	Housing Authority	yes
520	Libraries	
530	Commission for Historical Preservation	
540	Tax Sale Properties	
550	Docks and Wharves	
560	Housing and Community Development	yes
570	Market and Comfort Stations	
580	Other	
590	Other	
600	Office Buildings	
610	Parks and Recreation	
620	Police Stations	
630	Museums	
640	Fire Departments	
650	Public Works Properties	
660	Housing Authority	yes
670	Other	
680	Other	
690	Other	
700	Churches, Synagogues and Parsonages	
710	Church Schools	
720	Church Colleges	
730	Church Cemeteries	
740	Church Hospitals and Health Related Facilities	
750	Church Camps	
760	Other such as the Salvation Army or Missions	
770	Church Societies and Clubs	
780	Church-Aged and Rehabilitation Home	yes
790	Other	
794	Parking Lots	
795	Parking Lots	
800	Private Schools	
805	Payment in Lieu of Taxes	
810	Private Colleges	
820	Hospitals and Health Related Facilities	
830	Lodges	
840	Non-Profit Housing for the Elderly	yes
850	Boy Scout and Girl Scout Camps	
860	Other	
870	Other Camps	

880	YMCA or YWCA Camps	
890	Trade Associations	
900	Civic Organizations	
905	Community Owned Properties	
910	Clubs	
915	Research Organizations	
920	Historical Societies	
930	Museums	
940	Volunteer Fire Departments	
950	Fair Grounds	
960	Veterans Organizations	
970	Goodwill, Disabled Veterans Rehabilitation Centers and the Red Cross	yes
980	Private Cemeteries	
990	B & O Railroad	
991	B & A Railroad	
992	Conrail (Consolidated Railroad Corporation)	
993	National Railroad Passenger Corporation	
994	Penn-Central (Philadelphia-Washington-Baltimore Railroad)	

Appendix 3.2: Land use Categories

Land use category	Land use sub-categories
Agricultural and forest	Agricultural buildings Bare Ground Brush Cropland Deciduous forest Evergreen forest Feeding operations Mixed forest Open urban land Orchards/vineyards/horticulture Pasture Row and garden crops
Commercial	Commercial
Extractive	Extractive
Industrial	Industrial
Institutional	Institutional
Residential low-density	Low-density residential
Residential medium-density	Medium-density residential
Residential high-density	High-density residential
Water or wetland	Water Wetlands

Appendix 4.1: Comparison of predicted probabilities for multiple treatment analysis to obtain common support region

	TAZ distance from the closest Metro		
	< 0.5 miles	0.5 - 1 mile	> 1 mile
Probability of dose 0 (> 1 mile)			
mean	0.273	0.499	0.897
stad. dev	0.267	0.260	0.178
min	0.000	0.002	0.015
max	0.947	0.972	1.000
Probability of dose 1 (0.5 - 1 mile)			
mean	0.250	0.345	0.074
stad. Dev	0.187	0.175	0.125
min	0.007	0.019	0.000
max	0.770	0.853	0.784
Probability of dose 2 (<1 mile)			
mean	0.477	0.157	0.029
stad. Dev	0.331	0.184	0.072
min	0.020	0.004	0.000
max	0.993	0.922	0.800

Bounds for common support

Maximum	Pr(X)=0	Pr(X)=1	Pr(X)=2
dose 0	1.000	0.784	0.800
dose 1	0.972	0.853	0.922
dose 2	0.947	0.770	0.993
<i>Minimum of maxima</i>	<i>0.947</i>	<i>0.770</i>	<i>0.800</i>
Minimum			
dose 0	0.015	0.000	0.000
dose 1	0.002	0.019	0.004
dose 2	0.000	0.007	0.020
<i>Maximum of minima</i>	<i>0.015</i>	<i>0.019</i>	<i>0.020</i>

List of references

- Amin, K. & Capozza, D. R. (1993). "Sequential Development," *Journal of Urban Economics*, 24(1), 142-158.
- Badoe, D, & Miller, E.J. (2000). Transportation land-use interactions: empirical findings in North America, and their implications for modeling, *Transportation Research Part D: Transport and Environment*, 5(4), 235-263.
- Baum-Snow, N. & Kahn, M. (2000). The effects of new public projects to expand urban rail transit, *Journal of Public Economics*, 77, 241-263.
- Bayer, P., McMillan, R. & Rueben, K. S. (2004). What drives racial segregation? New evidence using Census microdata, *Journal of Urban Economics*, 56, 514-535.
- Behrman, J.R., Cheng, Y. & Todd, P.E. (2004). Evaluating preschool programs when length of exposure to the program varies: a nonparametric approach, *The Review of Economics and Statistics*, 86 (1), 108-132.
- Bento, A., Cropper, M., Mobarak, A.M., & Vinha, K. (2005). The impact of urban spatial structure on travel demand in the United States, *The Review of Economics and Statistics*, 87 (3), 466 – 478.
- Black, D., & Smith, J. (2004). How robust is the evidence on the effects of college quality? Evidence from matching, *Journal of Econometrics*, 121 (1-2), 99-124.
- Boarnet, M.G. & Crane, R. (2001). The influence of land use on travel behavior: specification and estimation strategies, *Transportation Research Part A: General*, 35, 823-245.

Bollinger, C. & Ihlanfeldt, K. (1997). The impact of rapid rail transit on economic development: The case of Atlanta's MARTA, *Journal of Urban Economics*, 42 (2), 179-204.

Bollinger, C. & Ihlanfeldt, K. (2003). The intraurban spatial distribution of employment: which government interventions make a difference, *Journal of Urban Economics*, 53 (3), 396-412.

Bowes, D. & Ihlanfeldt, K. (2001). Identifying the impact of rail transit stations on residential property values, *Journal of Urban Economics*, 50 (1), 1-25.

Byars, A., Greenwood, M., Greenwood, L. & Simpson, W. (2003). The effect of alternating steady-state walking technique on estimated VO₂max values of the Rockport Fitness Walking Test in college students, *Journal of Exercise Physiology* online, 6 (2).

Cameron, I., Kenworthy, J.R. & Lyons, T.J. (2003). Understanding and predicting private motorized urban mobility, *Transportation Research Part D: Transport and Environment*, 8 (4), 267-283.

Cervero, R. (1994). Rail transit and joint development – land market impacts in Washington, DC and Atlanta, *Journal of the American Planning Association*, 60 (1), 83-94.

Cervero, R. (2002). Built environments and mode choice: toward a normative framework, *Transportation Research Part D*, 7 (4), 265-284.

Cervero, R. (2004). Effects of light and commuter rail transit on land prices: Experiences in San Diego County, *Journal of the Transportation Research Forum*, 43 (1), 121-138.

Cervero, R. & Landis, J. (1993). Assessing the impacts of urban rail transit on local real estate markets using quasi-experimental comparisons, *Transportation Research Part A*, 27A (1), 13-22.

- Cervero, R. & Landis, J. (1995). Development impacts of urban transportation: A US perspective. In D. Banister (Ed.), *Transport and Urban Development* (pp. 136-156). London: E & FN Spon.
- Cervero, R. & Landis, J. (1997). Twenty years of the Bay Area Rapid Transit System: land use and development impacts, *Transportation Research Part A*, 31 (4), 309-333.
- Citizens' Advisory Committee. (1973). *Planning, zoning & development of central business districts & transit station areas*. Montgomery County, MD.
- Damm, D., Lerman, S.R., Lerner-Lam E., & Young, J. (1980). Response of urban real estate values in anticipation of the Washington Metro, *Journal of Transport Economics and Policy*, 14 (3), 315-336.
- Dehejia, R. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies, *The Review of Economics and Statistics*, 84 (1), 151-161.
- DiNardo, J. & Tobias, J.L. (2001). Nonparametric density and regression estimation, *The Journal of Economic Perspectives*, 15 (4), 11-28.
- Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation, *The American Statistician*, 37 (1), 36-48.
- Efron, B. & Tibshirani, R.J. (1998). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall
- Frölich, M. (2004a). Finite-sample properties of propensity-score matching and weighting estimators, *The Review of Economics and Statistics*, 86 (1), 77-90.
- Frölich, M. (2004b). Matching estimators and optimal bandwidth choice, Department of Economics, University of St. Gallen, version July 24.
- Frölich, M., Heshmati, A. & Lechner, M. (2004). A microeconomic evaluation of rehabilitation of long-term sickness in Sweden, *Journal of Applied Econometrics*, 19 (3), 375-396.

Gabriel, S.A. & Rosenthal, S.S. (1989). Household location and race: Estimates of a multinomial logit model, *The Review of Economics and Statistics*, 71 (2), 240-249.

Gatzlaff, D. & Smith, M. (1993). The impact of the Miami Metrorail on the value of residences near station locations, *Land Economics*, 69 (1), 54-66.

Grass, R.G. (1992). The Estimation of Residential Property Values around Transit Station Sites in Washington, D.C. *Journal of Economics and Finance*, 16 (2), 139-46.

Green, R. & James, D. (1993). *Rail transit station area development: small area modeling in Washington, DC*. Armonk, NY: M.E. Sharpe.

Hanna, W.J. et al. (1994). *Metro Stop? Metro, stop!: the politics of transportation planning*. College Park, MD: Urban Studies & Planning Program, the University of Maryland, College Park.

Härdle, W. (1990). *Applied nonparametric regression*, Cambridge, UK: Cambridge University Press.

Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from a job training programme, *Review of Economic Studies*, 64, 605-654.

Heckman, J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator, *Review of Economic Studies*, 65, 261-294.

Ihlanfeldt, K.R. (2003). Rail transit and neighborhood crime: The case of Atlanta, Georgia, *Southern Economic Journal*, 70 (2), 273-294.

Imbens, G.W. (2000). The role of the propensity score in estimating dose-response functions, *Biometrika*, 87 (3), 706-710.

Jalan, J. & Ravallion, M. (2003). Does piped water reduce diarrhea for children in rural India?, *Journal of Econometrics*, 112 (1), 153-173.

Joffe, M.M. & Rosenbaum, P.R. (1999). Propensity Scores, *American Journal of Epidemiology*, 150, 327-333.

Kain, J. (1964). A contribution to the urban transportation debate: An econometric model of urban residential and travel behavior, *The Review of Economics and Statistics*, 46 (1), 55-64.

Knaap, G.J., Ding, C., & Hopkins, L.D. (2001). Do plans matter? The effects of light rail plans on land values in station areas, *Journal of Planning Education and Research*, 21 (1), 32-39.

Lechner, M. (1999). Identification & estimation of causal effects of multiple treatments under the conditional independence assumption, IZA Discussion Paper No. 91. (Later published in M. Lechner & F. Pfeiffer (Eds.) *Econometric evaluations of active labor market policies in Europe*, Heidelberg: Physica/Springer, 2001)

Lechner, M. (2002). Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies, *The Review of Economics and Statistics*, 84 (2), 205-220.

Leuven, E. & Sianesi, B. (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing." Retrieved August 27, 2004 from <http://ideas.repec.org/c/boc/bocode/s432001.html>. Version 1.2.3.

Marchwinski, T. (1998). Economic impact of existing and new commuter rail service on retail and recreational spending in the vicinity of station areas, *Transportation Research Record*, 1623, 135-143.

Maryland-National Capital Park and Planning Commission. (1962). *On wedges and corridors: a general plan for the Maryland-Washington Regional District in Montgomery*

and Prince George's Counties. Silver Spring, MD Maryland-National Capital Park and Planning Commission.

Maryland-National Capital Park and Planning Commission. (1973). *Managing our future: An approach to urban growth policy for Montgomery County*, Urban Growth Policy Report no. 1, Silver Spring, MD: The Montgomery County Planning Board.

Maryland-National Capital Park and Planning Commission. (1977). *Prince George's County transit impact study: Summary report for the Addison Road corridor*. Riverdale, MD: Maryland-National Capital Park and Planning Commission.

Maryland-National Capital Park and Planning Commission. (2002). *Prince George's County Approved General Plan*. Upper Marlboro, MD: Maryland-National Capital Park and Planning Commission.

McDonald, J. & Osuji, C. (1995). The effect of anticipated transportation improvement on residential land values, *Regional Science and Urban Economics*, 25 (3), 261-278.

McDonald, J. & McMillen, D. (2000). Employment subcenters and subsequent real estate development in suburban Chicago, *Journal of Urban Economics*, 48 (1), 135-157.

McMillen, D.P. & McDonald, J.F. (2004). Reaction of house prices to a new rapid transit line: Chicago's Midway Line, *Real Estate Economics*, 32 (3), 463-486.

Metropolitan Washington Council of Governments (MWCOCG). (1979). *More than a subway: a chronicle of transit goals for the Nation's capital*, April.

Metropolitan Washington Council of Governments (MWCOCG). (1983). *Metrorail area planning: metrorail before-and-after study*, August.

Metropolitan Washington Council of Governments (MWCOCG). (2004). *COG/PB travel forecasting model, Version 2.1 D#50, User's Guide*, November 17. Retrieved September

12, 2005 from

<http://www.mwcog.org/transportation/activities/models/documentation.asp>.

Montgomery County Planning Board. (1969). "Update general plan (on wedges and corridors) for Montgomery County. Silver Spring, Maryland.

Moon, H. (1990). Land use around suburban transit stations, *Transportation*, 17 (1), 67-88.

Murin, W. (1971). *Mass Transit Policy Planning: An Incremental Approach*. Lexington, MA: DC Heath and Company.

O'Sullivan, S. & Morrall, J. (1996). "Walking distance to and from light-rail transit stations." *Transportation Research Record*, 1538, 19-26.

Pagan, A. & Ullah, A. (1999). *Nonparametric Econometrics*, Cambridge, UK: Cambridge University Press.

Pober, D.M., Freedson P.S., Kline, G.M., McInnis, K.J., & Rippe, J.M. (2002). Treadmill walk test to predict peak oxygen uptake in healthy adults ages 40 to 79 years, *Canadian Journal of Applied Physiology*, 27 (6), 575-588.

Rosenbaum, P.R. & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70 (1), 41-55.

Rubin, D.B. (1980). Bias reduction using Mahalanobis-metric matching, *Biometrics*, 36 (2), 293-298.

Ryan, S. (2005). The value of access to highways and light rail transit: evidence for industrial and office firms, *Urban Studies*, 42 (4), 751-764.

Saelens, B.E., Sallis, J.F. & Frank, L.D. (2003). "Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures." *Annals of Behavioral Medicine*, 25 (2), 80-91.

Smith, J., & Todd, P. (2004). Does matching overcome LaLonde's critique of nonexperimental estimators?, *Journal of Econometrics*, 125 (1-2), 305-353.

U.S. Census Bureau. (2001). *Traffic Analysis Zones*. Retrieved August 26, 2005 from http://www.census.gov/geo/www/cob/tz_metadata.html.

U.S. Census Bureau. (2005). Length of time for new residential construction. Retrieved September 2, 2005 from <http://www.census.gov/const/www/lengthoftimeindex.html>.

Voith, R. (1993). Changing capitalization of CBD-oriented transportation systems: Evidence from Philadelphia, 1970-1988, *Journal of Urban Economics*, 22, 361-376.