Demographics and Housing Choice.

Patrick Bajari and Matt Kahn.
Stanford and Columbia University.

Jan 3, 2000.
Preliminary- Comments Welcome.

bajari@leland.stanford.edu
kahn214@columbia.edu

Abstract

In this paper we study the relationship between household demographic characteristics and housing demand. We build a discrete choice model of housing demand that accounts for three key empirical problems: the fact that housing units and households are heterogenous, there are unobservable attributes to a housing unit and households with different places of work prefer to live locations closer to their job. We then use our estimates to study the causes of racial segration within the city of Philadelphia. We also provide estimates for how much households are willing to pay for a reduction in crime.

1 Introduction

For most U.S. families, housing comprises a significant percentage of annual expenditures and represents a large share of family wealth and savings. The total value in 1990 of U.S. single family homes was 5.4 trillion dollars or 61.7% of the total value of U.S. real estate (DiPasquale and Wheaton 1996). The choice of a house is inextricably bundled with a location, a community and the amount of time spent each day commuting to work. People spend a significant amount of time in their homes and peer groups for children are determined by housing choice. Many public policy issues are closely linked to housing choice such as racial segregation, commuting congestion, central city decline, savings and portfolio choice, peer group effects and child outcomes. Reliable estimates of household demand for tenure (whether to rent or own), structure, commuting and community are needed to evaluate the efficiency gains achievable through housing policies.

Hedonics and discrete choice methods have been the two leading approaches for identifying willingness to pay for housing attributes. Building on Rosen (1974), housing researchers have used hedonic techniques to estimate the marginal price of housing attributes. A few papers have attempted to estimate a "second stage" to identify structural demand parameters (Palmquist 1984, Gyourko and Voith 2000, Cheshire 1998). Discrete choice methods offer an alternative method for measuring demand parameters. Quigley (1985) estimates a nested logit model in which households first choose a community and then choose a housing structure in that community. Necheyba and Strauss (1998) study community choice in New Jersey to measure the demand for schools and Bayer (1999) estimates a discrete locational choice model to study school demand in California. This paper combines discrete choice models of differentiated products with hedonics to estimate structural housing demand parameters. We apply these tools to estimate housing demand in 1990 in the Philadelphia metropolitan area using micro data from the 1990 Census of Population and Housing and augmenting this with data on local public goods levels across the Philadelphia area. This study estimates household preferences over tenure status, commuting time, structure type and community.

The approach we follow in this paper resolves several important econometric problems faced in previous studies. First, the price of a housing unit will be correlated with unobserved (to the econometrician) attributes of the home. Standard discrete choice approaches (such as the well known conditional logit approach) which treat the housing price as uncorrelated with the error term will yield a downward biased estimate of the price elasticity of the demand for housing. Second, our methodology allows for a richer specification of the demand structure that will account for the interaction of demographic characteristics (size of household, income, age, and race) with characteristics of the housing unit (size, price, owner status and location) in the utility function. We present a detailed examination of the differences in housing demand by household type. Third, our methodology allows for an assessment of how households tradeoff commuting time versus community and structure type in choosing their favorite housing product. As we show below, the hedonic "two step" cannot measure the willingness to pay for inframarginal product attributes while the discrete choice approach can. Our estimates of how households tradeoff commute times versus other housing attributes provides new insights into the "wasteful commuting" puzzle (Small and Song 1992, Hamilton 1982 1991).

The empirical estimates of the structural model are used to measure a household's willingness to pay for ownership rather than renting, structure, community and commuting and to study how these parameters vary with with respect to household income, race age and household size. These parameters are then used to address two important questions in urban economics. Why does residential racial segregation persist? In a recent paper, Cutler, Glaeser and Vigdor (1999) argue that whites are willing to pay more to live near whites than blacks are willing to pay to live near whites. To test this "decentralized racism" hypothesis, we use our structural estimates of Philadelphia migrant willingness to pay to be in black communities by household type. We find that blacks demonstrate a slight revealed preference to live in communities featuring higher percentages of black households while white households do not exhibit strong preferences for living in non-black communities. Our explanation for why segregation persists focuses on place of work and the disutility from commuting. Blacks are twice as likely as whites to work in the center city. Since commuting lowers utility, this creates an incentive to live downtown which is re-enforced by black migrant households willingness to pay to live near other minority households.

The second urban issue we explore using our structural estimates is household willing to pay to avoid exposure to high crime areas. As murder rates fall across many major cities, New York City being the leading case, there is renewed optimism that center communities will experience a revitalization as "middle class flight" reverses (Mieskowski and Mills 1993). Berry-Cullen and Levitt (1999) report a large city to suburb migration in response to increases in crime. Our structural estimates complement their reduced form migration on crime regressions. Following Nevo's (1998) econometric strategy, we estimate how much households are willing to pay to avoid higher murder areas. Our estimates complement the labor literature's "value of life" estimates which are based on hedonic wage regressions which measure compensating differentials for occupational risk exposure (see Rosen and Thaler 1975, and Viscusi 1993). We find that households are willing to pay a great deal to avoid murder and surprisingly our estimates are roughly in line with Viscusi's (1993) survey estimates of the labor literature.

The Philadelphia housing market features over one million homes. In this market, there are too many products to estimate a discrete choice model where households choose among individual homes. Our estimation approach is to reduce the dimensionality of the problem such that each household chooses among 272 products. Thus we must assign each home in the Philadelphia metropolitan area to one of these "products". To simplify this paper's layout, we first outline what data we use. We present detailed summary statistics to provide a feel for how heterogeneous households are choosing different structures, communities and commuting times. We then carefully show how we construct the "products" that households choose among. We then present a structural model of housing choice. The estimates of this model are fully discussed to document population heterogeneity with respect to demand for housing attibutes. We then use our estimates to explore why racial segregation persists and willingness to pay to avoid crime.

2 Data

The raw data used in our empirical analysis comes from the 1990 Census of Population and Housing micro data (the 5% sample) for the Philadelphia metropolitan area. Philadelphia is an attractive area to study for several reasons. It is one of the largest metropolitan areas in the nation. The Philadelphia metropolitan area has an older housing stock than many parts of the country. Only 12.9 percent of the housing stock was built between the years of 1980 and 1990. The fact that most of the housing stock is over 10 years old allows us to view the stock of housing as largely predetermined. This greatly simplifies the econometric analysis because supply of housing should be modeled together simulataneously with demand in fast growing areas. Viewing the housing stock as exogenous is a less dangerous assumption in an area such as Philadelphia versus a sprawling area such as Phoenix or Las Vegas. Philadelphia is also an important area to study because there has been great concern about racial segregation in older northeast cities (Massey and Denton 1993). Philadelphia is large enough such that Census identifies many communities within the metropolitan area. In addition, employment is spread out across several employment centers such that, commuting times will differ (even for two identical households in the same community) because their place of work differs. As we discuss below, this variation in commutes is useful for studying community valuation.

The unit of analysis is a household with the demographic data based on the identified "head of household". Census data provides information on the housing structure a household lives in, its location, the location of where the head of household works, and basic demographic information on all members of the household. We use all of this information in estimating our housing demand models.

Our definition of community is determined by Census geographic identifiers called Public Use Micro Areas (PUMAs). There are 34 PUMAs within the Philadelphia MSA and 11 PUMAs in the center city. In most cases, PUMA boundaries were defined for the Census by State government. While PUMAs generally are aggregations of census tracts and urban places, they do not reflect the boundaries of political jurisdictions. PUMAs are intended to reflect "like" areas containing 100,000 people or more. Using the PUMA identifiers, we can aggregate the micro data to form any summary statistic such as a given PUMA's percent college graduates, or percent minority. The 1990 Census identifies place of work called "POWPUMAS". There are 14 POWPUMAs within the Philadelphia metropolitan area.

In Table One, we present summary statistics on the attributes of the stock of Philadelphia heads of households and recent migrants who have switched homes in the last five years. On average, recent migrants are 11 years younger than the stock and 21 percentage points less likely to own. Households sizes, income, commuting time, race, and sex are roughly the same across the two groups. Migrants are more likely to have children under age 18 in the household. Migrants represent 37% of the sample. There is significant heterogeneity with respect to commuting time. The average commuter spends 25.9 minutes commuting one way and the standard deviation is 16.9 minutes. Commuters who live near their job have an eleven minute shorter commute one way than commuters who do not live near their job. A commuter is defined to be a near mover if the first two digits of the PUMA identifier matches the POWPUMA identifier.

Our structural estimation will focus on the migrant sub-sample. To provide extra details about this group, in Table Two we report average consumption statistics stratified by household income levels. Household income is divided into three groups; poor (less than $25,000), middle class (between $25,000 and $50,000) and rich (greater than $50,000). Richer households are more likely to live in the suburbs (Margo 1992). The probability that a poor household lives in the center city is 49% while the probability that a rich household chooses a center city location is 16.8%. Rooms are a normal good. A rich household consumes 2.6 more rooms, and 1.3 more bedrooms on average than a poor household. Richer households are also more likely to live in newer structures. The average rich household lives in a community which is 12 percentage points less black, and which has 7 percentage points more college graduates than the average poor household. Ownership rises sharply with income such that a rich household has a 77% probability of owning while 23% of poor households own. Black household heads are over-represented among the poor but 6.8% of Philadelphia's rich migrants are black.

Table Three provides some additional facts about our Philadelphia migrant sample by estimating multivariate OLS regressions. In the left column of Table Three, we report a linear probability model of whether a household head is an owner as a function of household demographics. The omitted category is a white, single male. An extra person in the household raises the probability that a household owns by 3 percentage points. A female head of household is 2 percentage points less likely to own while a married man is 23 percentage points more likely to own. Both blacks and Hispanics are roughly 12 percentage points less likely to own. An extra $10,000 of income increases the probability of owning by 4 percentage points. The next two regressions study how migrant annual home price and rental expenditure varies across demographic groups. Married migrants spend $1115 more on owner occupied housing while married renters spend $205 more on housing than the omitted category. Blacks spend significantly less on housing each year regardless of tenure status. Note, that this is controlling for household income. A single black man who rents spends $1437 less per year than his white counter-part. An extra $10,000 of income increases expenditure on owner occupied housing by $855. Figure One provides additional details on household expenditure on housing as a function of income. For owners and renters separately, we have partitioned the sample by $1,000 dollar increments and caculated the household's expenditure share on housing (home price *.075/household income) and graphed this against household income. The graph shows that richer Phildalephia households spend a much smaller share of their income on housing than poorer households.

Returning to Table Three, the next regression studies how one way commuting time (measured in minutes) differs by demographic groups. Blacks have much longer commuting times, 5.9 minutes, which probably reflects greater public transit use. Richer households have longer commutes but the size of the coefficient, .24, is quantitatively small.

A straightforward method of classifying Philadelphia migrant households is whether they own or rent and whether they live in the center city or the suburbs. To document differences across households, in Table Four we report means by these four categories. The left column of Table Four reports means for center city owners. Center city owners spend almost 1/2 of what suburban owners spend annually on housing. The average center city owner lives in a community which features 10 percentage points less college graduates and 20 percentage points more blacks than the average suburban owner. Both suburban and center city owners feature households with .6 more people than the average renter household. Renters are much more likely to have a female head and much less likely to be married. Blacks are under-represented among suburban owners. Only 4.6% of suburban owners are black while 35.9% of city renters are black. The average city owner's income is just around $40,000 which is almost twice as high as center city renter's and 1/3 lower than suburban owners. Home owners in both the city and suburbs are more likely to have young children present in the household. Interestingly, city and suburban owners have roughly the same commuting time. This indicates that many suburbanites must work in the suburbs. Suburban renters have significantly shorter commutes than any other group. There are significant differences in the propensity to own a vehicle across the four categories. Just under 80% of center city owners and over 98% of suburban owners own a vehicle, while just 49% of center city renters own at least one vehicle. This suggests that center city residents will be more dependent on public transit and that their average commuting time to suburban employment centers could be quite high.

The first step in preparing to estimate our structural model is to estimate standard hedonic rental and home price OLS regressions. In the Census, rents and home prices are self reported as category variables. We take the midpoint of each category. There is very little top coding of the data in Philadelphia unlike other metropolitan areas such as Los Angeles. For example, only 2.4% of apartments and homes in Philadelphia are top coded. The highest rent is $1,000 a month and highest home price is $400,000. The dependent variable in the rental hedonic is annual rent for a given unit and the dependent variable in the home price regression is the home's reported price multiplied by 7.5% (see Blomquist, Berger and Hoehn 1988 and Gyourko and Tracy 1991 who also follow this convention). In these hedonic regressions, we control for structure attributes and community attributes. Structure attributes include; rooms, bedrooms, age of structure, whether the structure is a single detached unit or part of a complex, kitchen conditions, and dummies to control for when the household moved into the structure.

We run separate home price and rental regressions as a function of structure characteristics and PUMA level community characteristics. The regression therefore has the form:

$$\text{Housing Expenditure}=\text{Structure Variables}+\text{Community Variables}+\text{Error}$$

In the home price regression, the dependent variable is the household's annual expenditure on housing (home price multiplied by 7.5%) and in the renter regression, the dependent variable is the household's annual rent. The structure characteristics include; the type of structure (single detached, single attached, multi-unit dwelling etc.), the year the household moved into the structure, dummy variables to control for the age of the structure, the structure's number of rooms and bedrooms and dummy variables indicating whether the structure has a complete kitchen and toilet and whether it is heated using modern fuels (i.e. not wood or fuel oil). The home price regression (but not the rental regression) also includes a measure of annual housing upkeep. Using micro data from the 1995 American Housing Survey, we have constructed an index of annual housing upkeep expenditure as a function of household age and person characteristics. We use this equation to predict household maintenance by owners.

The hedonic estimates of the 1990 Philadelphia housing and rental regressions are presented in Table Five. The omitted category in these regressions is a single detached unit which was built after 1970 which is located in the suburbs and which uses older fuels and has an incomplete kitchen. The coefficients' units are annual 1989 dollars. For owners, the annual marginal cost of a room is \$395 and for renters it is \$374. An extra bedroom raises annual home prices by \$481 and rents by \$245. Most housing structures in Philadelphia are either single attached or single detached units. Owners pay \$3063 a year less for attached housing relative to detached housing while renters pay \$687 dollars less. There is a noticable difference in the discount for older owner occupied housing relative to older renter housing. The dummy variables for when the household moved into the structure convey important information. Unlike home owners, renters who have lived in their unit for a long time pay much lower rents for the quality adjusted unit. This suggests that such tenants have signed long term contracts. By controlling for year moved in, our hedonic prices the rental unit for a household currently considering such a unit.

Turning to the community attributes; homes in the center city cost \$738 more than their suburban counterparts while rentals are \$400 a year less. The community variables "PUMA % college graduate" and "PUMA % black" take on values from 0 to 1. The owner hedonic coefficients show that if a community has 10 percentage points more college graduates then home prices are \$1,269 higher and rents are \$525 higher. All else equal, if a community has 10 percentage points higher black population, home prices fall \$451 dollars and rents fall \$161 dollars. The college graduate variable proxies for high peer group and "role model" effects while a community's percent black proxies for local school quality and local crime levels.

In this paper we will carefully account for unobserved attributes of housing products. For example, Census data does not provide information on whether a household has a sauna or whether the household lives next to a beautiful golf course. As discussed in Epple (1987), the first stage of the hedonic "two step" only yields consistent implicit prices of product attributes if the unobserved attributes of the structure and the community its located in are uncorrelated with observables. To test the robustness of the structure estimates of the hedonic coefficient estimates presented in Table Five, we have dropped the community variables and included 34 PUMA fixed effects to control for any community level variables. In results that are available on request, we find that all of the structure attributes are robust to controlling for PUMA fixed effects.

The hedonic estimates of the structure attributes play a key role in our structural estimation. The Philadelphia housing market features millions of homes. There are "too many" products for households to choose from. To reduce the dimensionality of the problem, we aggregate similar homes into the same product. Intuitively, two homes located in the same PUMA which are both owner occupied which feature similar prices and similar structure attributes will be grouped as the "same" product. To aggregate these products, we need a measure of household expenditure on structure. Based on these hedonic estimates for renters and owners, we can predict household expenditure on structure by multiplying the hedonic regression coefficient on each structure attribute and multiplying this by a household's consumption of that attribute and then summing across the attributes. For example, if the price of a bedroom is $25 and the price of an extra room is $50, if these are the only structure attributes then a household who purchases a home with 2 bedrooms and 4 rooms would have a structure expenditure index equal to $250 per year. The construction of this structure index mirrors the approach described by Rotheberg et. al (1991), King (1976), and has been used extensively in the quality of life literature to rank communities by their local public goods levels (Blomquist, Berger and Hoehn (1988), Gyourko and Tracy (1991), Roback (1982), DiPasquale and Kahn (1999)).

To aggregate similar homes into a smaller set of "housing products". We create 8 products per PUMA using the structure index created from the hedonic regressions reported above. For each PUMA in Philadelphia, we calculate the 25th, 50th and 75th percentiles of the rental and home price distribution. For each puma we also calculate the structure index at the 25th, 50th and 75th percentiles. These are used as cutoff points to assign individual homes to "products".

We create 8 products per PUMA using the following approach. For a given household, we know whether they rent or own. Conditional on tenure status we know their annual housing expenditure and their predicted structure index consumption. This Table explains how we define products within a PUMA and what price we attach to the product.

Product	Renter	SIND>	PRICE>	Product Price
		Puma Median	Puma Median	
1	yes	yes	yes	75th quantile of
				PUMA rental pdf
2	yes	yes	no	25th quantile of
				PUMA home pdf
3	yes	no	yes	75th quantile of
				PUMA rental pdf
4	yes	no	no	25th quantile of
				PUMA home pdf
5	no	yes	yes	75th quantile of
				PUMA rental pdf
6	no	yes	no	25th quantile of
				PUMA home pdf
7	no	no	yes	75th quantile of
				PUMA rental pdf
8	no	no	no	25th quantile of
				PUMA home pdf

For example, in a given PUMA assume that a home is owned and that its predicted structure index is greater than the PUMA's median structure index and that the home's annual price is greater than the PUMA's median home price, then we assign this home to be product number 5 and assign this product a price equal to the 75th percentile of the PUMA home price distribution. Since there are 34 PUMAs in Philadelphia, this approach yields 272 products for households to choose between. Once we have assigned each home to a product, we create an identifier indicating which of the products the household chose. It is important to note that even though each housing unit may be "unique", given that we have aggregated products, many individual households have chosen the same "product".

The hedonic results presented above were based on all households in the Census. We use all of this information to construct the 272 housing products in the Philadelphia metropolitan area. In studying housing choice, we focus on residents of Philadelphia who have moved within the last 5 years. Conditional on the head of household's place of work, each household must simultaneously choose tenure status (to own or rent), the structural amenities of the unit (e.g. number of bedrooms, size, attached or detached) and the community in which the unit is located. Recent migrants are an attractive group to study because they have recently made a costly decision where they chose among the set of alternatives. We recognize that recent migrants may differ from the population as a whole (see Table One for differences with respect to observable demographics).

An important contribution of this research is to carefully model choice of commuting time as another element of location choice that each household must trade-off when comparing the benefits and costs of different housing products. For heads of households who work, we know their commute to work measured in minutes. We use this information to construct "product" commuting time. Taking a working head of household's place of work as given, we calculate what is the sample mean commute from every PUMA to that POWPUMA. Given that there are 14 POWPUMAS and 34 PUMAS, we calculate 476 means. A commuter who works in POWPUMA 2600 will recognize that a PUMA near POWPUMA 2600 offers a shorter commute to his job than other PUMAS while a commuter who works in POWPUMA 5200 will recognize that the same residence features a long commute for him. If all employment was centralized at one central business district location, then average commuting times would be the same for all people who choose the same PUMA. Since employment is not fully centralized, two heads of households who are considering the same product in a given PUMA will face different commuting times from that PUMA if they work in different POWPUMAS. If a person does not work, we estimate the average commute time in the whole PUMA which represents a measure of access to the Central Business District. Table Six provides some information on the possible variation in commute times which households face. In each row of this Table, we report the minimum, median, and maximum one way average commute time (in minutes) to each POWPUMA. For example, for workers who work in POWPUMA 2900, there is no PUMA with an average commute time less than 17.93 minutes to work, while the median PUMA's commute to this POWPUMA is 42.6 minutes. As shown in Table One, the average commute is less than 26 minutes. This indicates that households are taking into account commuting time in choosing where to locate. Our structural model will provide new insights into the "wasteful commuting" literature which has studied why households do not live right next to their job. Our structural model will focus on the tradeoffs heterogeneous households who differ by demographic type face when choosing among heterogenous homes.

Commute times differ greatly depending on whether a household lives and works in the center city or whether the household lives and works in the suburbs or if the household commutes between the city and suburbs. The average commute for a household which lives in the city and works in the city is 25.84 while if this household works in the suburbs its average commute is 32.73 minutes. A commuter who lives in the suburbs and works in the city faces the longest average commute of 38.03 while a suburban resident who also works in the suburbs has the shortest average commute of 22.43 minutes.

In this study, we take a household head's choice of where to work as exogenous. Not all demographic groups are equally represented across employment centers. Black households are much more likely to live in the center city and to work in the center city than non-black households. Table Seven reports three cross-tabs for the whole migrant sample, for white migrants and for black migrants. For each household, we create a dummy variable which equals one of if the household lives in the center city and which equals zero if the household lives in the suburbs and another dummy which equals one if the household head works in the center city and which equals zero if the household head works in the suburbs. The top tabulation shows that roughly 22% of the total sample live and work in the center city, while over 58% of the sample live and work in the suburbs. Almost three times as many household heads live in the suburbs and commute to the city than vice versa. Over 50% of black migrants live and work in the center city while 30% of the black migrants live and work in the suburbs. 64% of white migrants live and work in the Philadelphia suburbs. Whites are much more likely than blacks to live in the suburbs and commute to the city for work. An equal number of black migrants live in the city and commute to the suburbs and vice-versa.

.To summarize this section, we have used the Philadelphia hedonic price regressions to create a structure index for each household. Using this information and information on which PUMA each migrant household chose, we assign each household to having picked one of 272 housing products. These products differ with respect to their community attributes, structure index, and price. In addition, commuters will take into account their commute time if they choose a given product. Using data on place of work, we know what a commuters' commute would be for every possible product he chould choose. Census data's rich demographic information allows us to model product choice allowing for population heterogeneity with respect to demand for product attributes.

Given that we now have explained what data is used in the structural analysis, it is important to compare our "product" construction to more standard IO research such as Petrin, or Berry, Levinson, and Pakes (1995). In the IO literature, consumers take the product attributes as given. For example, General Motors determines the engine horsepower of a Buick.[1] In our urban problem, a community's attributes are an emergent property of the set of households who choose to live within its borders. If all college graduates demand to live in a given PUMA, then this PUMA will feature high levels of "mba". In our analysis we assume that migrants take as given such community attributes as "percent college graduate" and "percent black". Some may be concerned that there is a "reflection" problem (Manski 1992). It is important to note that the

[1] It is true that in a Veblen style "bandwaggons" model, the average consumer attributes would be associated with the vehicle. For example, Mercedes is aware that if it lowers its price too much, its vehicle may lose "cache" because it will no longer be associated with being a signal of a household's prosperity. Becker and Murphy (1992) explore fads with respect to restaurant pricing as restaurants take into account that part of their product's attributes are the average characteristics of its consumers.

community attributes are based on the attributes of all households, not just migrants, within the PUMA. Given that only 37% of Philadelphia households are migrants (see Table One), it is not an outlandish assumption to assume away that households change a community's local public goods vector by their choice of community. We could even further reduce this problem if we focused on the subset of migrants who lived outside the Philadelphia MSA five years before. We recognize that migrants change the community if the marginal migrant's attributes differ from the community average. It is true that in a more "fluid" metropolitan area, such as Las Vegas, or other fast growing sprawl capitals, migrants represent a much larger fraction of the population in this case, there could be mulitiple equilibria such that households who want to live in college educated areas may move simultanenously without being sure where are the high "mba" areas.

3 A Model of Housing Demand.

In this section, we build a model of housing demand for the city of Philadelphia. Our econometric modeling strategy is motivated by three fundamental empirical concerns. The first is that both housing units and households are heterogeneous in important ways. In our analysis, we take account of two dimensions along which housing products may be differentiated. The first is the physical amenities of the housing product such as the number of rooms, bedrooms and the number of square feet. Second, houses differ by location. Neighborhoods have important externalities such as crime, the quality of schools and the distance to urban amenities such as shopping or cultural activities.

In our model, we allow households' valuation of the characteristics of a housing unit to depend on demographics. For example, large families, all else held constant, will value extra bedrooms and space more than smaller families. Also, younger households may place a premium on the freedom to move with ease and therefore value owning less than older households. In many previous studies of housing demand, such as King (1976) and Follain and Jiminez (1985), researchers have modeled housing as a continuous good and used hedonic weights to aggregate housing characteristics.

Aggregating housing characteristics into a single index destroys important information about consumer's preferences and limits the researcher's ability to explain what types of households match to what types of housing units.

We are by no means the first to incorporate these concerns into our econometric framework, but we attempt to do so in a unique way. In more traditional, hedonic methods, such as Epple (1987), researchers attempt to recover the willingness to pay for various attributes of a home. However, a key problem for hedonics is that many of the attributes of a housing unit may be unobserved to the econometrician. In our econometric analysis, we address this problem by estimating a large number of product level fixed effects-one for each of the 272 housing units.

Recent empirical work in applied microeconomics has emphasized that in a discrete choice model, prices will be correlated with unobserved product attributes for the simple reason that higher quality commands a higher price in the market place. Failure to account for these product level unobservables will cause the researcher to underestimate own price elasticities. This has been documented in recent empirical work by Petrin (1999) in his study of demand for minivans, Nevo (1999) in his study of demand for breakfast cereals and in Berry, Levinsohn and Pakes (1995) in their study of automobile demand. In all three cases, these researchers find that price elasticities are underestimated by an order of magnitude when the econometrician fails to account for unobserved product level heterogeneity[2]. In our reduced form analysis of the previous section, we find that the observed product characteristics can account for only between 40 to 60 percent of the observed variation in prices. Clearly, products have a number of attributes that are not present in the data set and we attempt to control for this by using fixed effects within our discrete choice model.

A key problem for hedonic models of housing demand, as discussed in Epple (1987), is sorting on unobservables. In our analysis, we attempt to address this problem by allowing for an interaction between a product level fixed effect and a household's demographic characteristics. The

[2] Not surprisingly, in earlier models of housing demand, we found that the households' sentivitiy to prices was an order of maginitude lower than the specification studied in this paper.

fixed effect represents attributes of the housing unit not included in our analysis such as crime, school quality and the unobserved structural amenities. Since the household's utility includes demographics interact with unobserved housing attributes, the approach taken here models the sorting of observed demographics on unobserved housing attributes and allows us to recover the household's valuation for both observed and unobserved attributes of the housing unit.

A final way in which demographics enter into the choice of housing units is through the budget constraint. Not all housing units are affordable to all households. For instance, mortgage companies will typically limit the maximum value that a family can spend on housing as a percentage of income in determining eligibility for loans. In our analysis, we do not allow households to purchase housing units that are more than 50 percent of a family's pre-tax income. This cut-off point is roughly in line with practices by mortgage companies as well as is appears to be a reasonable cut-off point for a budget constraint since this represents the 90^{th} percentile of the empirical distribution of housing expenditure as a percentage of income. If we fail to include a budget constraint in the analysis, we will misspecify consumer's choice sets and make false inferences about their underlying tastes. While admittedly, this is a rough approximation to consumer's true choice sets, this set is typically ignored in much applied work using discrete choice models.

Lastly, we believe that commuting is an important factor in the choice of housing unit. Our reduced form analysis indicates, not surprisingly, that households prefer locations that are close to the place of work for the head of household. In our discrete choice model, each head of household computes her commute time from alternative PUMAs, taking the place of work as given. Therefore, we can estimate willingness to commute using census data. To the best of our knowledge, ours is the first study that incorporates commuting data this way into a discrete choice model of housing demand. A hedonic "two step" approach to measuring willingness to pay to avoid commuting is only likely to be effective in a monocentric world where 100% of the employment is located in the center of the city. If employment is dispersed across many employment centers, then distance to any one employment center may not be capitalized into the first stage hedonic home price regression (like the one presented in Table Four) and the hedonic researcher would have no hope of recovering willingness to pay in the "second stage" of the Rosen "two step" (Rosen 1974).[3]

3.1 An Econometric Model.

In this section, we describe a discrete choice model of housing demand. The primitives of the model are household preferences, demographics and product characteristics. The econometrician is assumed to observe both individual purchase decisions and demographic traits. The conditional indirect utility function of a consumer is a function of the observed product attributes, household demographics and models parameters. There are $i = 1, ..., I$ households and $j = 1, ..., J$ housing products products. Formally, we write a household's utility function as $U(x_j, \xi_j, p_j, d_i, \varepsilon_{ij}; \theta)$. The vector x_j is a $k \times 1$ vector of observed characteristics of product j and ξ_j is a vector of characteristics that are unobserved to the econometrician, but observed by the households. The price of product j is p_j, the individual's demographic characteristics are d_i which are assumed to be observable to the econometrician, ε_{ij} is a disturbance to the consumer's decision making that is drawn independently for each i and j and θ is a vector of parameters.

The data used in estimating the model are:

x_j : Observable characteristics of product 1.

- The price of product j, p_j.
- The unit's structure index measured in dollars, $sind_j$.
- The percentage of head of households who are black in the PUMA associated with the j^{th} product, $mblack_j$.
- The percentage of head of households who are college educated in the PUMA associated with the jth product, mba_j .
- An indicator variable for whether the product is owned or rented, own_j.

[3] The hedonic approach does not allow two observationally identical households to value the same product differently. Yet, if one head of household works near that product while another head of household works far from that product, the former household will gain greater utility from living their than the latter.

d_i The demographic characteristics of household i.

- An indicator variable for whether or not the head of household is white, $white_i$.
- The number of people in household i, $person_i$.
- The income of household i, $income_i$.
- The age of the head of household i, age_i.

c_{ij} The commute time of household i to product j, taking the household's place of work as given.

The utility function used in this research is of the form:

$$\begin{aligned} u_{ij} \; = \; & \xi_j + \beta_1 \log((income_i - price_j)/10,000) + \beta_2 \log((sind + 85.0)/10,000) \\ & + \beta_3 \log(mblack_j) + \beta_4 \log(mba_j) + \beta_5 own_j + \pi_1 \log((income_i - \\ & price_j)/10,000) * white_i + \pi_2 * \log((income_i - price_j)/10,000) * \\ & person_i + \pi_3 * \log((income_i - price_j)/10,000) * \log(income_i/ \\ & 10,000) + \pi_4 \log((income_i - price_j)/10,000) * \log(age_i) + \pi_5 * \\ & \log((sind + 85)/10,000)) * white_i + \pi_6 * \log((sind + 85)/10,000)) \\ & * person_i + \pi_7 * \log((sind + 85)/10,000)) * \log(income_i/10,000) + \\ & \pi_8 * \log((sind + 85)/10,000)) * \log(age_i) + \pi_9 * \log(mblack_i) * \\ & white_i + \pi_{10} * \log(mblack_i) * person_i + \pi_{11} * \log(mblack_i) * \\ & \log(income_i/10,000) + \pi_{12} * \log(mblackj) * \log(age_i) + \pi_{13} * \\ & \log(mba_i) * white_i + \pi_{14} * \log(mba_i) * person_i + \pi_{15} * \log(mba_i) * \\ & \log(income_i/10,000) + \pi_{16} * \log(mba_i) * \log(age_i) + \pi_{17} * own_j * \\ & white_i + \pi_{18} * own_j * person_i + \pi_{19} * own_j * \log(income_i/10,000) \\ & + \pi_{20} * own_j * \log(age_j) + \pi_{21} * \xi_j * white_i + \pi_{22} * \xi_j * person_i + \end{aligned}$$

$$\pi_{23} * \xi_j * \log(income_i/10,000) + \pi_{24} * \xi_j * \log(age_i) + \pi_{25} * \log(c_{ij})$$
$$* \log(income_i) + \varepsilon_{ij}$$

Two assumptions about the utility function bear discussion. First, several of the terms, such as price, income and the structure index were divided through by 10,000. This is essentially a scaling assumption that makes the likelihood function well-behaved numerically. Second, we added 85 to the structure index to make sure that we did not take the log of a negative number since the value of $sind$ was negative for some housing products.

In our discrete choice model, we incorporate a simple version of a budget constraint. Household i was assumed only to be able to choose those products j for which the ratio $\frac{p_j}{income_i} < .5$[4]. As we argued previously, it is important to incorporate a budget constraint into the analysis because not all housing goods are affordable for all households. Let $J(i)$ denote the set of products that are affordable for household i.

For identification purposes we normalized $\xi_1 = 0$ and we make an assumption that ε_{ij} comes from a double exponential distribution with variance of 1. Define $\widehat{u}_{ij}(x_j, \xi_j, p_j, d_i, \varepsilon_{ij}; \theta) = u_{ij}(x_j, \xi_j, p_j, d_i, \varepsilon_{ij}; \theta) - u_{i1}(x_j, \xi_j, p_j, d_i, \varepsilon_{ij}; \theta)$. Let $I(i,j)$ be an indicator variable for the event that household i chooses product j. The probability that household i chooses product j is $P(I(i,j) = 1|x, \zeta, p, d_i))$ is then:

$$P(I(i,j)) = \frac{\exp(\widehat{u}_{ij}(x_j, \xi_j, p_j, d_i, \varepsilon_{ij}; \theta))}{\sum_{k \in J(i)} \exp(\widehat{u}_{ik}(x_k, \xi_k, p_k, d_i, \varepsilon_{ij}; \theta))}$$

To form the full likelihood function, we will also incorporate the census weights associated with each household into the analysis. Let cen_i be the census weight associated with household i. Let $L(I; x, \zeta, p, d, cen)$ be the likelihood function for the observed choices, that is, I is the vector of all observed choices, x is the vector of observed product characteristics for all products,

[4] Those data points in the sample where this equality was violated were removed. Furthermore, all households with income reported at less than \$8,000 were removed. This caused only a small reduction in sample size, from 2376 to 2031 reporting households.

ζ is the vector of all product level fixed effects, p is a vector of all prices for all products and d is the vector of all household level demographics. The likelihood function then satisfies:

$$L(I; x, \zeta, p, d_i, cen) = \prod_i P(I(i,j))^{cen_i}$$

The model parameters are done using maximum likelihood estimation.[5]

3.2 Discussion.

Our econometric framework is an extension of a standard logit model where we allow for interactions between housing, community and commute time with household level demographics. Unlike the logit model, however, the restrictive substitution patterns implied by the independence of irrelevant alternatives do not hold in the aggregate for our model. This is because we include demographic interactions in our model. In our econometric model, a home is a combination of 6 attributes: a value of $sind$, mba, $mblack$, own and ξ as well as $price$. Unlike some previous studies which aggregate housing consumption into a single index, our model allows for five dimensions along which homes may differ. Since there is a full set of demographic interactions, our econometric model will allow us to explore how different demographic groups match to heterogenous housing units.

Our econometric framework attempts to deal with product level unobservables by including fixed effects in the estimation. For each product j there is a separate ξ_j that captures utility from the unobservable attributes of the product. As we mentioned earlier, the observed attributes of a home are only capable of explaining 40 to 60 percent of the variation in price. Therefore it is clear that unobservables are present. Since price is correlated with attributes of the home that are not observed to the econometrician, failure to include fixed effects would result in a serious endogeneity problem.

[5] The estimation algorithm was coded by the researchers in Fortran. The researchers used the IMSL library's numerical optimization proceedures to find the parameter values used. The sources code is available from the authors upon request.

In the econometric model, there is also a full set of demographic interactions with ξ_j. Therefore, we are able to estimate household's willingness to pay for both observed and unobserved housing attributes. Since different values of the demographic variables lead to different marginal valuations of the unobservables, our framework allows us to account for how observed demographic groups sort on unobservables. In a hedonic specification, as Epple (1987) explains, if households sort on unobservables, the estimates produced in a two step hedonic procedure will not be consistent. The specification of the household's maximization problem here allows us to address this problem at least partially. Our analysis does assume, however, that unobserved household attributes are independent of ξ_j.

After allowing for a full set of fixed effects and demographic interactions, our model contains 301 parameters to estimate the demand for 272 products. While it is possible to argue for a more flexible specification by using a richer error structure or random coefficients as in Petrin (1999), Nevo (1999) and Berry, Levinsohn and Pakes (1995), this may well prove to be computationally infeasible. Because of the wealth of demographic and commuting data and the fact that there are 272 products, we found that it takes roughly 2 to 3 days to compute our maximum likelihood estimates using a Sun Workstation. Our econometric framework was therefore motivated by what we felt were the first order problems in correctly estimating housing demand. We leave extensions of the framework to future researchers who will undoubtedly have access to more powerful computational tools.

It is important to note that we have forced the Philadelphia migrants to choose among the 272 products offered in the Philadelphia metropolitan area. Clearly, if the price of all of these products skyrocketed then many of these households would move to other cities such as Chicago. In our econometric model, we choose not to include an outside good. The drawback of this choice is that even if the prices of all products become arbitrarily large, consumers will still live somewhere in Philadelphia. This is clearly false. However, in previous versions of this model, we found clearly implausible results when we included an outside good into the analysis, such as wrong signs on various coefficients and clearly implausible elasticities. This should not be surprising since there is no single "good" that characterizes what heterogeneous households will consume if they move out of Philadelphia. The option chosen by the richest and poorest household will clearly differ. However, this is not allowed by an assumption of an outside good.

A limitation of our model is that households face a static maximization problem. Housing is a durable good and moving is costly therefore the consumer must evaluate her trade-offs intertemporally. To properly form a structural model of the dynamics of decision making, panel data on individual level choices is typically required. However, this type of data is not available in the census. Data sources such as the PSID do contain panels on individual households, however, we would loose the rich description of the variation of the housing stock within a city and how different demographic groups choose to locate within a city.[6] Fortunately, as Rust (1994) points out, the type of model we estimate can be correctly interpreted as the value function of a household's dynamic program (See Rust (1994) for a more complete discussion.). All of our demographic variables would be natural state variables in a dynamic model of housing and location choice. For example, current income and race are important indicators of future income. Higher current income today one would expect to increase the expectation of future income and, if there is discrimination in labor markets, when the indicator variables white is zero, the expectation of future income falls. These considerations are important when interpreting the results from our analysis.

Consider, for instance, a positive coefficient on the interaction between $white$ and own (as it is in our estimates). This could reflect higher marginal valuation of home ownership in each period among whites. An alternative explanation, however, is that whites have higher expectations for future earnings due to discrimination in labor markets and, in anticipation of higher future earnings, whites spend more on home ownership than non-whites. It is not clear how we can

[6] By studying housing choice within a single local labor market, we sidestep the difficult problem of how households choose which local labor market to locate in and then conditional on this what is their best housing product to consume.

resolve these issues without panel data on households. Since our demographic variables would be natural state variables in almost any dynamic model of housing consumption, to interpret the results we must bear in mind that households face not only static trade-offs, but also make their decisions in light of expectations about the future.

4 Parameter Estimates and Results.

The parameter estimates for the model are as follows:

Parameter Estimates

Parameter	MLE	S.E.	Parameter	M.L.E.	S.E.
β_1	17.72		π_1	-3.6670	
β_2	0.8229		π_2	-1.4673	
β_3	1.3515		π_3	8.3505	
β_4	3.8692		π_4	-2.8733	
β_5	-3.8172		π_5	0.5750	
π_6	1.9526		π_{16}	-0.5710	
π_7	0.9067		π_{17}	0.6055	
π_8	-0.2093		π_{18}	0.2101	
π_9	-0.6872		π_{19}	1.0760	
π_{10}	-2.6861		π_{20}	0.3757	
π_{11}	8.4141		π_{21}	0.4157	
π_{12}	-0.1918		π_{22}	-0.10206	
π_{13}	-0.4422		π_{23}	0.6172	
π_{14}	-0.5465		π_{24}	-0.1273	
π_{15}	0.4299		π_{25}	-1.6628	

Next, we report a set of results about how an average member of the population would value a ten percent change in any one of the 5 variables that describe the observable characteristics of the product. To see how these computations work, first of all describe a method for measuring how a household values a 10 percent increase in the variable mba. We consider two possible consumption bundles, in the value of mba in the first bundle is mba_1 and the value of the in the second bundle is $mba_2 = 1.1 * mba_1$. The total amount of non-housing consumption in the first consumption bundle is $c_1 = income_1 - price_1$ and the total amount of non-housing consumption in the second bundle is $c_2 = income_2 - price_2$. If all other characteristics of the housing bundles are constant across the first and second bundle the consumer will be indifferent between the first and second bundle if and only if:

$$\widetilde{\beta}\log(c_1/10,000) + \widetilde{\pi}\log(mba_1) = \widetilde{\beta}\log(c_2/10,000) + \widetilde{\pi}\log(mba_2)$$

where:

$$\begin{aligned}
\widetilde{\beta} &= \beta_1 + \pi_1 white_i + \pi_2 person_i + \pi_3 \log(income_i/10,000) + \pi_4 \log(age_i) \\
\widetilde{\pi} &= \beta_4 + \pi_{13} white_i + \pi_{14} person_i + \pi_{15} \log(income_i/10,000) + \pi_{16} \log(age_i)
\end{aligned}$$

By a simple computation it follows immediately that:

$$\frac{c_1}{c_2} = (1.1)^{\widetilde{\pi}/\widetilde{\beta}}$$

We will, in the text, refer to $\frac{c_1}{c_2}$ as the consumption ratio. Similar willingness to pay calculations can be made for other product characteristics.

We use our structural estimates to present a set of willingness to pay calucations when the demographic characteristics are set equal to their population averages. An average household in our sample has a value of white equal to 0.8340, the number of persons equal to 2.735, an average income of 48,361 and an age of 38.1511. The average amount of expenditure on nonhousing consumption is 34,642 dollars (pretax). The results are best interpreted in the center of our sample, so we include a sample mean and standard deviation for all variables. In the following table we summarize these willingness to pay measures and evaluate them at the center of our sample. We increase the value of $sind$, $mblack$, mba by ten percent and compute the consumption ratio as in the above expression. We also change the value of the indicator variable own from zero to one and double the commuting time for the household. The willingness to pay calculation is found by computing the difference between c_1 and c_2 for a household with consumption of 34,642 dollars (pre-tax) on nonhousing goods.

Variable	Average	S.D.	Consumption Ratio	Willingness to Pay
$sind$	5489.00	2351.57	1.0145	502.30
$mblack$	.1478	.1870	1.001	34.642
mba	.1365	.1209	1.0043	148.96
own	.49	.499	1.0248	859.12
$commute$	25.93	16.9	0.872	4503.00

Since this table is computed at the sample average, it masks the fact that willingness to pay varies greatly within the population. In housing markets, heterogenous households match to heterogenous housing units. In order to understand this matching process, it is useful to study how willingness to pay measures vary as the demographic characteristics of the household vary. This in turn will help us to understand the process of how different households match to different housing units in the city.

In Figure 2, we report how the willingness to pay for structure varies with respect to household demographics. We double the amount of the variable $sind$, at the sample average, an increase from \$2,351 to \$4,702. All household demographic characteristics are set to the population average. with the number of persons equal to 2.735, income is 48,361 and an age of 38.1511. To compute the willingness to pay for a doubling of structure, we use equations 4.1 through 4.3. In the first figure, when we vary the variable white from zero to one, the willingness to pay increases to \$4,300 while blacks are willing to pay \$2,400. There is a black/white differences in willingness to pay for structure even holding income fixed.[7] This could be due to a number of factors. As we discussed above, the indirect utility function we estimate is correctly interpreted as a reduced form for a household's value function. A household's race is a state variable that contains important information about the probability of events in the future, such as future earnings profiles, for instance. Therefore, in interpreting these results it is important to keep in mind that the higher willlingness to pay by whites for structure could be do to life cycle

[7] Note that structure's units are dollars. Using the hedonic regressions presented in Table Five, one can measure what is the willingness to pay for any one component of the structure index such as the willingness to pay for an extra room

factors such as higher expected permanent income. Another interesting result from figure two is the number of people in the household has a large impact on the demand for structure.

While richer households spend more on structure than poorer households, the willingness to pay for an increase in structure falls as income rises. This is consistent with our reduced form results that suggest that the expenditure share on housing tends to have a convex shape. As households become richer, they tend to spend a lower overall percentage of income on housing services. This finding has important implications for Alonso, Muth, Mills (AMM) model of urban form. As discussed in Wheaton (1977), if the rich have a high willingness to pay for housing structure which is greater than their willingness to pay to avoid commuting, then optimal housing choice provides an explanation for why we observe the wealthy in the suburbs (where structure is cheaper to purchase but commutes are longer) rather than in the center city. Our findings are in accord with Wheaton (1977) that the income elasticity of demand for structure is too low to support the AMM explanation for why the wealthy live in the suburbs. decentralization.

In Figure 3, we study willingness to pay to live in minority communities as a community's percent black "mblack" is doubled from .1478 in our sample to 0.2956. It is possible that minority households are eager to live in the same community with other households of similar backgrounds. Such communities may offer greater access to social networks and basic child care services. White households may fear minority communities and be willing to pay not to live there even if good housing near jobs are available. A household's willingness to pay for this attribute is relatively insensitive to household demographics. Our results indicate that non-whites have a slight preference for living close to other non-whites. This small estimate is surprising. In this paper, we have only assumed that households face a budget contraint. Racism is certainly another constraint which would limit the products which a non-white household would have to choose from. If blacks could only choose among a handful of products which are concentrated in black communities, then our structural approach would yield estimates that black households greatly desire living in black communities when in fact they were *forced* to live in these communities (Kain and Quigley 1975).

A white household is roughly indifferent, when all other demographic variables are set near their average, between living in a PUMA with a value of .1478 for $mblack$ and a PUMA with a value of 0.2956 for $mblack$. This, perhaps surprising, result needs to be explained. Our structural model controls for the PUMA's percent college graduates and for a fixed effect. Thus, a community's percent black is not proxying for unobserved community attributes such as poor schools or neighborhood crime. These product level attributes are captured by the product fixed effect. We believe that if we did not control for these attributes, then white households would exhibit a much greater disutility from living near non-whites.[8] While non-whites reveal a slight preference for living with non-whites, our results appear to stand in contrast with a prominent recent paper by Cutler, Glaeser and Vigdor (1999) who argue that racial segregation persists across cities due to "decentralized racism". This hypothesis states that whites are willing to pay more than blacks to live with other whites and that this contributes to segregation. Our estimates of the Philadelphia migrants' utility function parameters indicate that other factors, namely the pursuit of ownership, living near college graduates, and commute minimization dwarf the willingness to pay to avoid minorities. Below, we will return to this important urban question of why blacks are over-represented in the center city.

A community's percent college graduate, mba, is an important local public good. As discussed by Case and Katz (1991) and Rauch (1993), there are likely to be greater intellectual spillovers in more educated communities where children are more likely to have access to "role models". For a doubling of the sample average from .1365 to 0.273, the average household would be willing to pay $148. Non-whites are willing to pay more for "role models" than whites. Minority

[8] In a recent study, DiPasquale and Kahn (1999) present a hedonic home price regression for the Los Angeles county in 1990. Exploiting the intra-county variation across this area's 58 PUMAs, they show that after controlling for the structure's attributes and the PUMA's crime level, public school quality, and human capital proxies, that the PUMAs racial composition is barely capitalized into home prices. Dropping these objective local public goods indicators leads to much larger negative values of % minority on home prices.

households are more likely to have an unmarried female head. A household attempting to increase child quality will recognize that there is a production function where community inputs are substitutes for parental inputs. If a parent is not living in the household, then a better community may be needed to offset the lost parental inputs. As the number of people in the household increases, the willingness to pay to live near college graduates falls. Keep in mind that our measure of income is household income. In a larger household, per-capita income is lower. Since high quality neighbors are a normal good, poorer families are likely to afford living near them especially if they need to purchase more structure. Richer households also are willing to pay more for a doubling of mba. A household with an income of $17,000 is willing to pay $750 while a richer household with an income of $50,000 is willing to pay $1,250. Also, all other things held constant, when the head of household is older, living in a community with a higher percentage of college graduates becomes less important. This is additional evidence that highly educated communities are viewed as an input in child quality production.

Ownership is always discussed as a key component of the "American Dream". A famous paper by Mankiw and Weil (1989) argued that shifting demographics, namely the aging dissaving Baby Boomers, would lead to a 30% crash in home prices. Demographic changes can only influence home prices if different demographic groups have very different preferences for owning versus renting. For example, if younger or smaller households do have low willingness to pay to own, then this would represent evidence in favor of the Mankiw and Weil (1989) hypothesis. While our model is static and saving for a downpayment to purchase a home is an important real world phenomena (see Engelhardt and Mayer 1996), our structural estimates provide new insights into the willingness to tradeoff consumption for ownership. Whites are willing to pay more for houseing than non-whites. Once again, it is important to remember that demographic variables are state variables. This difference could be attributed to differences in future expected earnings between demographic groups. Homeownership increases in importance, not surprisingly, as the number of members in the households increases. As the size of the household increases from 1 to 7, a household with seven persons and all other variables set at their demographic means, would be willing to sacrifice up to$13,856.80 per year to move from the status of a renter to an owner. Ownership becomes more important as income increases. Lastly, as the age of the head of household increases, ownership becomes more important. It is important to remember that our sample consists solely of migrants so we our estimation results are not picking up older households "locked in" to their owner occupied home and thus never moving.

Lastly, we consider the willingess to pay to commute less time. There is certainly popular concern about the growth in congestion in urban areas as vehicle use rises (Downs 1991). Public transit is well known to be a time intensive transportation technology (LeRoy and Stonselie 1983). The possible efficiency gains achievable through congestion pricing crucially depend on how much households are willing to pay to avoid commuting longer. Center city politicians also have a stake in measuring willingness to pay to avoid commuting. As documented above, center city workers who live in the center city commute an average of 1/2 an hour less per day than center city workers who live in the suburbs. If households have a great disutility for commuting, then small improvements in central city schools and reduced crime could have a large effect on reducing middle class flight to the suburbs (Mieskowski and Mills 1993). At the sample average, this implies that a household is willing to sacrifice $6,045 dollars to avoid a doubling of one way commute time, from 25 to 50 minutes. If an household works 260 days per year, a twenty-five minute increase in one-way commuting time implies an extra 216 hours on the road each year. Our estimates imply that a household values its time at 28 dollars per hour at the sample average. While we expected to find that the willingness to pay to avoid commuting rises with income (with an elasticity close to one), we actually find that commuting is not very income elastic. As recently discussed by Calfee and Whinston (1998), new products such as high quality vehicles with entertainment systems may actually give the rich a comparative advantage in commuting relative to poorer households. To summarize our findings on the AMM hypothesis, we find that both the income elasticity of demand for structure and commuting are quite low.

The willingness to pay analysis has focused on a simple comparative statics holding all demographic factors at their means. In a heterogeneous metropolitan area like Philadelphia, the "average person" may not exist. To study the willingness to pay for "real people", we create 24 household types and compute for each of these households their willingness to pay for structure, ownership, community and commuting. There are two age categories (household heads ages 30 or 45), two race categories (white and non-white), two household sizes (2 people or 4 people) and three household income levels. The household income levels ($18,700, $34253, and $54,000) represent the 25th percentile, median and 75th percentile of the Philadelphia migrant household income distribution.

Table Eight provides estimates of willingness to pay for 24 demographic types labeled type 1-24. There are 3 income groups representing low income, median income, and high income (25th, median and 75th percentile of the empirical distribution, respectively), 2 racial groups (whites and non-whites), 2 age groups (age 30 and age 45), our sample are migrants who are younger and two household sizes (2 people and 4 people). This is a useful way to display population variation in willingness to pay for structure, commuting, ownership and commuting. Our thought experiment is to measure in annual dollars willingness to pay for a doubling of each attribute, holding all other attributes at their sample means and holding demographics at the levels indicated in each row. Each of the 24 rows represents a different demographic group and each of the right five columns represents a willingness to pay for a different attribute. To keep the discussion intelligible, the left column of Table Eight identifies which demographic "type" we are discussing.

Controlling for income, age and household size, poor whites are consistently willing to pay a multiple of over 3 times more for structure than are poor non-whites. Larger white families are willing to pay more for structure than smaller white families. For example, comparing type #8 to type #6, a poor white household of 4 where the head is age 45 is willing to pay $5510 per year for a doubling of structure while if the same household has only people it would be willing to pay $2208 for the same increase in structure. Middle class households are presented in types #9-#16. Perhaps surprisingly, willingness to pay for structure does not rise with income. Comparing type #16 to type #8, controls for age, race and household size, the richer white family is actually willing to pay less for structure than the poorer white family ($4953 to $5510). Middle class non-white households are willing to pay more for structure than poor non-white households of equal size and equal age (compare type #7 and type #15). Increasing a household's income from the median of the empirical distribution to the 75th percentile, increases both white and non-white household willingness to pay for structure but the slopes are not large. A comparison of types #16 and type #24 shows that increasing a white four person household's income by $20,000 increases annual willingness to pay for structure by $1700 and if that household were non-white would increases its willingness to pay for structure by $1400 (see types #15 and type #23).

The next column of Table Eight shows willingness to pay to live in a community where its percent black has doubled from 13% to 26% black. As discussed above, our thought experiment holds all other local public goods constant and these estimates reflect the fact that product fixed effects have been included in the estimation. Non-white households, across all demographic groups, are willing to pay to live in a community with more black households. The minimum willingness to pay (see Type #1) is $638 per year for this increase while the maximum willingness to pay is $1,462 (see Type #19). Perhaps surprisingly, while whites are willing to pay less to live in communities with greater percentages of blacks, the willingness to pay across the 12 white demographic groups are only negative for 6 and the most negative value is only $-438 per year for poor, white older households with 4 people in the household. Note that unlike this household, rich white young households of only 2 people are willing to pay to live near blacks (see type #18).

We interpret this finding that controlling for local public goods provision (such as crime and school quality), race in of itself has a small positive effect on attracting black migrants and only a small effect on deflecting white migrants to choose other products. While qualitatively in accord with Cutler, Glaeser and Vigdor's "decentralized racism" hypothesis, our findings suggest that other factors may be more important in explaining white/black separatiion.

There is significant heterogeneity across household types with respect to willingness to pay to for a 13 percentage point increase in one's PUMA percentage of college graduates. Holding other demographics constant, non-whites are always willing to bid more than whites for college educated role models. Larger households are willing to bid less than smaller households for more college graduates and in fact, larger households often have a negative willingness to pay for % college educated. We interpret this finding as evidence of a budget constraint. Smaller households are willing to pay a considerable amount for access to communities filled with college graduates. Both young, rich, small minority and white households are willing to pay $2,663 per year (see type #17) and $2353 per year (see type #18), respectively.

Unlike some of the other attributes, the 24 household types exhibit less variation in willingness to pay to own. Not surprisingly, the key demographic driver is household income. Poor households, regardless of age, race or household size, do not want to own. Keep in mind what this means. Households recognize that the opportunity cost of renting is owning which features a higher price and lower consumption of non-market goods. Given our structural estimates of beta 1 and pi 1, we know that poorer households place a greater value on non-housing consumption. Thus at the given prices of the owner occupied products, poor households do not want to purchase these products. Among middle class households, young, small families are not willing to pay to own. For these households, Whites are willing to pay more to own than non-whites. This may reflect permanent income differentials across races. The willingness to pay for the opportunity to own is quite low even for middle class large households. Note that household type #16 is willing to pay $1,442 per year to own. Rich white households are willing to pay over 4 times more than rich non-white households to own ($3669 to $826) in comparing types #24 to type #23. We recognize that ownership is a function of household expected income stream and household asset accumulation. It is well known that whites have higher assets than non-whites and this may explain the static "preference" differential.

Commuting is the final attribute that we measure variation in willing to pay for. The thought experiment is to see how much different households would be willing to pay to avoid a doubling of their daily commute. Dividing these numbers by 216 hours provides a measure of how different demographic groups value their time. Across the 24 groups, the median is $15.1 per hour, and the 10th percentile of the value of time distribution is -$5.74 and the 90th percentile is $-26.98, with a mean of $-16.42. These summary statistics are intuitively plausible and the variation is reasonable. Thus our explanation for observed wasteful commuting is simple. Heterogeneous households value heterogenous structures differentially. While they are willing to pay not to commute, prefered homes are often not located next to the household's place of work. Households reveal tradeoffs that appear quite reasonable. Only in a world with homogenous preferences and a homogenous housing stock, would we be surprised if we observed households not living as close as possible to their jobs. In this paper we have documented the extent the population heterogeneity with respect to housing attributes.

5 Applying Our Estimates to Two Fundamental Urban Issues

5.1 Why Are There So Few Black Households Living in Suburbia?

In our 1990 Philadephia migrant sample, 60% of whites and 30% of non-whites live in the suburbs. Living in the suburbs is a normal good (Margo 1992). Controlling for income and household size, a white migrant household has a 33 percentage point higher propensity to live in the suburbs than a black household and a 28 percentage point higher probability of working in the suburbs. Non-white households are twice as likely as white households to work in the center city (see Table Seven). Conditional on working in the center city, commute minimizers will search for a housing structure in the center city. Given that blacks have a slight preference for living with other blacks this re-enforces the process. Blacks live in the center city because they are commute minimizers who work in the center city and are willing to pay a slight bit to live near other blacks.

This explanation does not explain why blacks are twice as likely to work in the center city than whites. We have implicitly assumed in this analysis that households have found a job and then search for a housing unit conditional on where they work. Spatial mismatch theorists (see Kain 1992) might hypothesize that we have reversed the true process. Due to racial segregation, minority households search for the best housing unit within a predefined set of inner city ghetto areas and then upon finding a housing unit choose a commute minimizing job. Since they are forced to live in the center city, commute minimizers who may also not know about suburban jobs or be able to access them, settle on taking a more accessible center city job. It is important to note that we are studying a sample of recent migrants who tend to be younger, and a majority of these minority households do own at least one vehicle. In 1990 Philadelphia there are a large number of middle class and rich blacks who have greater choice across the products than the poorer households or middle class minority households in earlier decades. In this paper, we have assumed that minorities face the same constraints as whites (namely the budget constraint that households cannot purchase products whose price is greater than 50% of their income).

We now conduct two counterfactual experiments to gain some further insight into why non-whites do not live in the suburbs. According to our structural model, it is utility maximizing for 67.06 percent of non-whites to live in the city. One possible explanation for why this is true is that it is simply too expensive for non-whites to afford homes in the city. To test this, let's replace the income of all non-whites in the data set with the sample median income of 48,461 dollars. With this replacement of income, now 67.95 percent of non-whites choose to live in the city, an increase. If income is increased to 60,000 the proportion of non-whites who live in the city is 64.98 percent and if the increase in income is to 90,000 dollars then 59.05 percent of non-whites will live in the city. Clearly, the estimates from our structure model imply that it takes an increases in income to 90,000 dollars to move a mere seven percent of non-whites from the city to the suburbs holding all else fixed. According to the parameter estimates of in our model, income alone cannot explain why non-whites are overrepresented in the center city relative to the suburbs.

A second hypothesis about why non-whites live in the center city is that they disproportionately work in the center city. We now run a counterfactual experiment where we replace all non-whites place of work to the POWPUMA that is farthest from the center city and compute the utility maximizing choice for each household and keep all other demographic variables the same. In this experiment, 49.85 percent of non-white households choose to live in the city, a nearly twenty percent decrease. Our model seems to indicate that segregation in the city is deeply tied to the place of work rather than income. Large changes in income do not result in a migration of non-whites from the center city but a change in place of work will incent non-whites to change location.

5.2 How Much Do Households Value Crime?

The structural model yields 272 estimates of the product unobserved attributes. As discussed in Nevo (1998), these estimates contain valuable information which we use in a "second stage" OLS regression. To provide some intuition, consider a simple example from labor economics. Suppose that a researcher estimated a standard wage regression where the dependent variable is a person's wage and explanatory variables include the person's age, education and occupation. Suppose that the researcher knew what state each person lives in. In this case, the researcher can include state fixed effects in order to get a "clean" estimate of the returns to being in any given occupation. This approach yields 50 state fixed effects which can be used in a second stage regression. In this second stage regression (which would have 50 observations), the researcher would fit a state's wage fixed effect as a function of state attributes. For example, states with high taxes may have to pay higher wages as a compensating differential or states with good climate are likely to pay lower wages for the same job (Gyourko and Tracy 1991, Roback 1982).

Using the 88 fixed effects from the center city PUMAs (11 PUMAs and 8 products per PUMA), we follow the exact same approach. In our case, we are interested in whether differences across the 11 PUMAs can explain the variation in the fixed effects. Up to this point we have only used 1990 Census of Population and Housing micro data, but any location specific data can be used in this second stage as long as it can be merged to the PUMA data. In a separate study, Cummings, DiPasquale and Kahn (2000) have collected data on murders per-capita in Philadelphia in the mid-1990s by census tract. After geocoding the street address of each murder, they located in which census tract the murder took place.[9] We aggregate this census tract data up to the PUMA level and use this variable as a "second stage" regressor. A second explanator regressor is the PUMA's share of owner occupied housing. As discussed by DiPasquale and Glaeser (1999) and Green and White (1995), there is some empirical evidence of the local benefits of living in a community with more home owners because they have a greater stake in the health of the community and are more likely to have "planted roots" in the community. A third variable included in the second stage regression is a PUMA's distance (measured in miles) to the Philadelphia Cityhall. Proximity to the center city and its amenities are culture are likely to be an amenity. We have also tried to include a measure of local school quality across the center city PUMAs. We found that this variable was highly negatively correlated with the murder rate.

Table Nine reports two second stage regressions. We focus on the left column's results which includes 87 observations based on the 88 products in the center city. One outlier is driving the results. In the right column we report the estimates based on the 88 observations but we focus on the results which drop this outlier which are presented in the left columns. The first point is that PUMAs with higher murder rates have a lower product fixed effect, while PUMAs with higher ownership rates and greater proximity to Cityhall have higher product fixed effect values. The dependent variable's units are measured in utility (see the utility function in equation 4.1),

[9] Since they only collected data on murder rates within the center city's 11 PUMAs, we set the murder rate equal to zero in the suburbs. In results that are available on request, we have estimated this second stage regressions only for the 88 center city products and continue to find qualitatively similar results.

thus we calculate willingness to pay for changes in murder exposure and PUMA ownership. To judge the economic magnitude of the OLS estimates effect, we return to our structural estimates in particular pi21-pi25. These are the demographic interactions which tell us how different demographic groups value a change in the unobserved product fixed effect. Our second stage regression allows us to ask; if the murder rate reduces the product fixed effect, and if the rich value a high product fixed effect, then how much are the rich willing to pay to avoid a higher murder community?

Reduced form migration regressions of out-migration regressed on center city crime yield a large estimate of flight (Berry Cullen and Levitt 1999). While researchers have estimated hedonic regressions of rents on local crime levels, most risk capitalization studies have focused on labor market valuation of risk. Viscusi's (1993) comprehensive review mostly of labor compensating differentials for risk estimates a value of life of $5 million. This estimate is based on hedonic earnings regressions. There has been concern that labor market estimates under-estimate the compensating differential for risk exposure because safety is a normal good (Garen 1988). Housing market may be a better market for measuring value of life than the labor market.

Table Ten reports our estimates of how the 24 demographic groups value avoiding high murder communities and living in communities with greater levels of "stakeholders". The thought experiment is to measure how much a household is willing to pay to avoid exposure to a PUMA with 3 per 10,000 people were killed in 1994 and 1995 on average versus in a PUMA with no murder. While all demographic groups are willing to pay to avoid this chance to be murdered, whites are willing to pay more than non-whites, larger families are willing to pay more than smaller families and the rich are willing to pay more than the poor. A poor black, young, small family (type #1) must be paid $1,395 per year to avoid this risk. Dividing $1,395 by .0003 yields a value of life for the household of $2.3 million (pre-tax) dollars. Rich white, large households where the head is age 45 (see type #24) are willing to pay over $8100 to avoid this risk per year and this translates into a revealed value of life of $7.3 million dollars per person. These large but plausible estimates suggest that a permanent decline in center city crime will have a large effect on attracting households to communities which they would not have considered in the past.

6 Conclusion

This paper has shown how hedonic and discrete choice techniques can be used in tandem to estimate demand parameters for ownership, structure, community and commuting. Micro census data allows us to estimate a rich demand system documenting the population heterogeneity in demand for these atttributes. Hedonic estimation was needed to reduce the dimensionality of the product space. Even in a single meteropolitan area there are millions of homes. To shrink this set down to a computationally feasible set of "products", we estimated standard hedonic home price regressions. Home prices were decomposed into that part due to structural attributes and that part due to community capitalization. Using the structural attribute price estimates and knowing which community each housing unit is located in, we partitioned all housing units into falling into 272 mutually exclusive and exhaustive categories. Without hedonic methods, we could not aggregate "like" homes into the same bins. After constructing the 272 Philadelphia products, we estimated discrete choice model which differs from more standard approaches. Product fixed effects along with demographic interactions were included.

The estimated coefficients were used to provide new insights into how households tradeoff various housing attributes. While our study features many new findings concerning willingness to pay for housing, we believe that the major insights of our paper concern black suburbanization propensities, household willingness to pay to avoid risk and resolving the "wasteful commuting" paradox.

This research can be extended in a number of directions. This paper has focused solely on measuring willingness to pay for structure, community, tenure and commuting without analyzing how product choice affects the products themselves. Household locational choice changes communities if the marginal entrant differs from the community average. Our approach offers the opportunity to return to Schelling's question on how and when communities "tip". Future research could embed our structural model of housing choice and community into a "Schelling" tipping model of community formation and racial composition. For example, if a community features a high percentage of black households and a high percentage of college graduates and it is close to a major employment center, will the typical white household be willing to pay to live there or be willing to pay not to live there? Our approach yields the necessary parameters to predict whether there would be "white flight" from such a community and whether a black household would be most likely to win the bidding to replace this household. In a more realistic world featuring a heterogeneous housing stock and employment locations, our structural parameters are crucial parameters for determining community racial stability.

[1] Bayer, Patrick. An Empirical Model of Local Markets for Elementary Education, Residential Location, School Choice, Education Production and Competition. 1999 Stanford University Working Paper.

[2] Berry, Steven; James Levinsohn, and Ariel Pakes. Automobile Prices in Market Equilibrium. Econometrica;63(4), July 1995, pages 841-90.

[3] Berry, Steven. Estimating Discrete Choice Models of Product Differentiation" Rand Journal of Economics;25(2), Summer 1994, pages 242-62.

[4] Blomquist, G., M. Berger, and J Hoehn, 1988. New Estimates of Quality of Life in Urban Areas. American Economic Review 78: 89-107.

[5] Calfee, J. and C. Winston, The Value of Automobile Travel Time: Implications for Congestion Policy, Journal of Public Economics, 69(1), 83-102 (1998)

[6] Case, Anne, and Larry Katz, "The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Youth," NBER Working Paper No. 3705, 1991.

[7] Chesire

[8] Cullen, Julie-Berry and Steve Levitt. Crime, Urban Flight, and the Consequences for Cities. Review of Economics and Statistics, February 1999

[9] Cummings, Jean, Denise DiPasquale and Matthew Kahn. "The Renter to Owner Transition in Philadelphia" City Research mimeo. 2000.

[10] Cutler, David, and Edward Glaeser, "Are Ghettos Good or Bad?" Quarterly Journal of Economics, 1997, pp. 827-871.

[11] Cutler, David., Edward Glaeser, and Jacob Vigdor, The Rise and Decline of the American Ghetto. Journal of Political Economics. June 1999

[12] DiPasquale, Denise and Edward Glaeser. Incentives and Social Capital: Are Homeowners Better Citizens? Journal of Urban Economics, 45, 1999. 354-384.

[13] DiPasquale Denise and Matthew Kahn "Measuring Neighborhood Investments: An Examination of Community Choice by Migrants" forthcoming in Real Estate Economics

[14] DiPasquale, Denise; Wheaton, William. Urban economics and real estate markets Englewood Cliffs, N.J.: Simon and Schuster, Prentice Hall, 1996, pages xii, 378.

[15] Downs, Anthony, Stuck in Traffic. Brookings Institution 1991.

[16] Epple, Dennis. Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. Journal of Political Economy;95(1), February 1987, pages 59-80..

[17] Follain, James and Emmanuel Jimenez. "Estimating the Demand for Housing Characteristics: A Survey", Regional Science and Urban Economics, 15, 1, 1985, 77-107.

[18] Garen, John . Compensating Wage Differentials and the Endogeneity of Job Riskiness. Review-of-Economics-and-Statistics;70(1), February 1988, pages 9-16..

[19] Green, Richard and Michelle White. Measuring the Benefits of Homeownership's Effect on Children. Journal of Urban Economics, 41, 1997. 441-461.

[20] Gyourko, Joe. and Joe Tracy. 1991. The Structure of Local Public Finance and the Quality of Life. Journal of Political Economy. 91(4): 774-806.

[21] Gyourko, Joe and Richard Voith. Brookings-Wharton Paper Series.

[22] Hamilton, Bruce. Wasteful Commuting Again. Journal of Political Economy;97(6), December 1989, pages 1497-1504..

[23] Hamilton, Bruce. Wasteful Commuting. Journal of Political Economy;90(5), October 1982, pages 1035-51..

[24] Kain, John F. (1992) "The Spatial Mismatch Hypothesis: Three Decades Later," Housing Policy Debate; v3 n2, pp. 371-460.

[25] Kain, John F.; and John Quigley. Housing Market Discrimination, Homeownership, and Savings Behavior. American Economic Review;62(3), June 1972, pages 263-77..

[26] Kain, John, and John Quigley. Housing Markets and Racial Discrimination. New York: NBER, 1975.

[27] King, A. Thomas, "The Demand for Housing: Lancastrian Approach," Southern Economic Journal, 43, 2, 1976, pp 1077-1087.

[28] Mankiw, Greg. and David Weil. 1989 The Baby Boom, Baby Bust and the Housing Market. Regional Science and Urban Economics. (May) V. 19. p. 235-258.

[29] Manski, Charles, "The Reflection Problem". Review of Economic Studies 1993.

[30] Margo, Robert. "Explaining the Postwar Suburbanization of the Population in the United States; the Role of Income, Journal of Urban Economics 31, 301-310 1992.

[31] Mieskowski, Peter and Edwin Mills. "The Causes of Metropolitan Suburbanization". Journal of Economc Perspectives V7; n 3 135-147. 1993

[32] Massey, Doug. and Nancy. Denton. 1993. American Apartheid: Segregation and the Making of the Underclass. Cambridge: Harvard University Press.

[33] Nechebya, Thomas and Robert Strauss. "Community Choice and Local Public Services: A Discrete Choice Approach," Regional Science and Urban Economics 28(1), 51-74 (1998)

[34] Nevo, Aviv. A Research Assistant's Guide to Random Coefficient Discrete Choice Models of Demand. 1998 University of California at Berkeley Working Paper.

[35] Palmquist, Raymond. Estimating the Demand for the Characteristics of Housing. Review of Economics and Statistics;66(3), August 1984, pages 394-404.

[36] Petrin,

[37] Quigley, John. Consumer Choice of Dwelling, Neighborhood, and Public Services. Regional Science and Urban Economics;15(1), February 1985, pages 41-63.

[38] Quigley, John. Nonlinear Budget Constraints and Consumer Demand: An Application to Public Programs for Residential Housing. Journal of Urban Economics;12(2), September 1982, pages 177-201.

[39] Rauch, James. 1993. Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. Journal of Urban Economics. 34(3): 380-400.

[40] Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." Journal of Political Economy 82 (January/February 1974): 34-55.

[41] Rosen, Sherwin and Richard Thaler. "Value of Life". 1975 NBER

[42] Rothenberg, Jerome et-al. The maze of urban housing markets: Theory, evidence, and policy Chicago and London: University of Chicago Press, 1991, pages viii, 549.

[43] Rust, John.

[44] Small, Kenneth; Shunfeng Song,-Shunfeng "Wasteful" Commuting: A Resolution. Journal of Political Economy;100(4), August 1992, pages 888-98.

[45] Viscusi, Kip. "Value of Life" Journal of Economic Literature 1993.

[46] Wheaton, William. Income and Urban Residence: An Analysis of Consumer Demand for Location. American Economic Review;67(4), Sept. 1977, pages 620-31.

Table One

Summary Statistics for the 1990 Philadelphia Sample

	all heads of households		all households who have switched homes in the last 5 years	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Annual home price	8804.205	6230.252	10267.160	6590.392
Annual rent	5303.591	2371.698	5721.597	2301.919
migrant in last 5 years	0.371	0.483	1.000	0.000
persons	2.645	1.505	2.599	1.484
female	0.365	0.482	0.363	0.481
married	0.537	0.499	0.494	0.500
black	0.181	0.385	0.165	0.371
Hispanic	0.025	0.155	0.038	0.190
owner	0.700	0.458	0.486	0.500
household income	4.184	3.715	4.136	3.559
age	50.090	17.350	39.711	15.058
kids present	0.346	0.476	0.412	0.492
central city	0.356	0.479	0.326	0.469
one way commute	25.801	16.910	25.930	16.918
% live near job	0.582	0.493	0.563	0.496
one way commute conditional on live near job	21.235	15.188	21.041	14.955
one way commute conditional on not live near job	32.162	17.131	32.234	17.215
Observations	65333		23769	

In this table, “migrant in the last 5 years”, “female”, “married”, “black”, “Hispanic”, “owner”, “kids present”, “central city” are dummy variables. Commute times are measured in minutes and prices are in 1989 dollars.

Table Two

Descriptive Migrant Attributes

	Household Income		
	Less than $25000	greater than $25000 less than $50000	greater than $50000
live in center city	0.490	0.290	0.168
rooms	4.155	5.263	6.707
bedrooms	1.834	2.386	3.117
home built before 1960	0.586	0.492	0.367
PUMA % black	0.214	0.130	0.089
PUMA % college graduate	0.286	0.310	0.355
household is an owner	0.232	0.507	0.775
household head is black	0.277	0.132	0.068
observations	7442	8721	7606

PUMAs are Census defined geographical units. There are 34 PUMAs within the Philadelphia metropolitan rea.

Table Three

Descriptive Regressions for Migrants

	Dependent variable						
	own	travel time	live near job	PUMA % college	PUMA % black	Annual Owner housing expenditure	Annual Renter housing expenditure
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
persons	0.031	0.165	0.005	-0.012	0.004	253.935	36.010
female	-0.022	-0.164	0.040	-0.000	0.003	-325.597	5.311
married	0.227	1.816	-0.064	-0.011	-0.030	1114.780	205.421
black	-0.120	5.846	0.076	-0.036	0.253	-4502.979	-1436.543
Hispanic	-0.126	-0.546	0.112	-0.067	0.084	-3926.164	-1079.734
Income ;10,000)	0.036	0.238	-0.014	0.009	-0.006	854.960	423.309
age	0.003	0.004	0.000	-0.000	-0.000	49.257	2.230
constant	0.041	22.515	0.614	0.335	0.137	2586.913	4692.647
observations	23769	18215	18215	23769	23769	12934	10835
R2	.24	.02	.03	.11	.32	.42	.30

The omitted category is a male, single, white household head. Race, sex, and marital status are dummy ariables. Only OLS coefficients are reported. Migrants are household who switched homes in the last five ears.

Table Four

Philadelphia Migrant Attributes Stratified by Tenure and Locational Choice

	center city owners	center city renters	suburban owners	suburban renters
annual housing expenditure for owners	5317.897		11791.100	
annual housing expenditure for renters		5002.422		6224.812
% of PUMA who are college graduates	0.243	0.297	0.340	0.322
% of PUMA who are black	0.271	0.322	0.066	0.080
persons in household	2.868	2.288	2.986	2.240
female head of household	0.361	0.531	0.208	0.436
married	0.546	0.242	0.740	0.347
black	0.267	0.359	0.046	0.136
Hispanic	0.059	0.066	0.013	0.039
household income ($10,000s)	4.023	2.250	6.039	3.156
age of head of household	42.024	37.939	40.529	39.069
children under age 18	0.452	0.337	0.516	0.322
one way travel time to work	27.553	26.971	27.215	22.950
% of households who do not own a vehicle	0.209	0.509	0.019	0.158
observations	2290	3360	10644	7465

Table Five

Housing Hedonic Regressions

	owner		renter	
	Coef.	Std. Err.	Coef.	Std. Err.
single unit attached	-3062.921	249.706	-687.003	130.537
2 units	-36.555	285.632	-187.480	117.251
3-4 units	218.622	475.030	-140.375	115.038
5-9 units	-2951.519	543.845	241.997	121.952
10-19	-2677.025	382.486	381.499	129.933
20-49	-1403.825	758.418	607.777	150.081
50+ units	15.775	921.130	1251.810	257.715
Moved in 1985-88	-30.874	82.283	-475.027	54.688
Moved in 1980-84	-290.686	93.929	-1131.072	100.166
Moved in 1970-79	138.414	98.107	-1284.145	146.728
Moved in 1960-69	-260.651	113.992	-1498.423	195.751
Moved in pre-1959	246.884	130.012	-1555.743	133.882
Built in 1960-1980	-4888.475	370.134	-231.926	160.485
Built pre-1960	-6629.833	373.865	-655.332	210.605
rooms	395.923	57.290	374.495	50.667
bedrooms	481.286	128.671	245.342	77.617
complete kitchen	473.688	330.742	276.388	199.992
modern fuel	-328.213	217.039	269.674	56.400
central city	738.142	284.411	-400.017	193.173
PUMA % college	12694.410	1463.296	5249.073	650.457
PUMA % black	-4501.678	584.435	-1610.001	631.691
maintenance index	1.501	0.147		
constant	2952.964	933.545	2637.300	358.955
observations/R2	51879	.59	17545	.36

Levels specification where the dependent variable is annual housing expenditure as a function of structure attributes and community tributes. The omitted category is a single detached unit built after 1980 in the suburbs.

Table Six
Philadelphia Commuting Times By Place of Work

place of work	minimum average commute PUMA	median average commute PUMA	maximum average commute PUMA
2600	17.64372	35.3329	58.22554
2700	17.04254	38.89571	60
2800	16.47355	34.39015	54.48186
2900	17.93318	42.64948	72.70718
3000	17.29977	36.82274	61.80488
4500	6	29.59877	60
4600	14.55909	31.88525	76.11111
4700	5	36.00167	69.16666
4800	15.31415	34.41718	72.1875
4900	10	25	50
5000	5	25.06744	60
5100	14.77263	32.83773	68.61111
5200	12.76013	31.2	60
5300	5	32.82803	90

The Census data identifies 14 place of work PUMAs (POWPUMAs) and 34 residential PUMAs. For ach POWPUMA, we calculate the average commute time to that POWPUMA from each of the 34 residental PUMAs. In each row of this Table, we report the minimum, median, and maximum one way average commute time (in minutes) to each POWPUMA. or example, for workers who work in POWPUMA 2900, there is no PUMA with an average commute time less than 17.93 minutes to work.

Table Seven

Migrant Place of Work and Place of Residence

All Migrants

Live in city	Work in city: 0	1	Total
0	270967	62339	333306
1	24762	103362	128124
Total	295729	165701	461430

White Migrants

Live in city	Work in City: 0	1	Total
0	240892	52606	293498
1	17051	63331	80382
Total	257943	115937	373880

Black Migrants

Live in city	Work in City: 0	1	Total
0	18491	6611	25102
1	5867	30914	36781
Total	24358	37525	61883

Table Eight

Philadelphia Migrant Willingness To Pay by Demographic Group

Demographic Group						Willingness to Pay in $ per year for a doubling of;				
type	white	person	age	income	income - price	structure	PUMA % black	PUMA % college graduate	probability of ownership 0% to 100%	commute time 25 min to 50 min
1	0	2	30	18700	13200	656.781	638.468	1022.85	-1736.93	-897.625
2	1	2	30	18700	13200	1926.56	15.0889	951.91	-1582.21	-1371.31
3	0	4	30	18700	13200	982.006	831.751	12.4123	-1727.03	-1240.61
4	1	4	30	18700	13200	3807.41	-107.278	-1042.62	-1435.66	-2371.83
5	0	2	45	18700	13200	653.335	640.181	908.112	-1752.08	-1008.3
6	1	2	45	18700	13200	2208.33	-112.666	745.889	-1574.45	-1647.33
7	0	4	45	18700	13200	1039.29	871.277	-325.933	-1747.55	-1462.34
8	1	4	45	18700	13200	5510.44	-438.156	-2237.71	-1348.86	-3334.85
9	0	2	30	34253	27562	1627.82	951.455	1756.12	-1394.07	-2441.64
10	1	2	30	34253	27562	3196.96	101.6409	1554.98	-443.033	-3165.63
11	0	4	30	34253	27562	2092.98	1095.64	420.599	-820.555	-2988.7
12	1	4	30	34253	27562	4456.94	17.6774	-375.611	744.797	-4149.66
13	0	2	45	34253	27562	1644.39	922.482	1573.04	-1223.38	-2632.99
14	1	2	45	34253	27562	3400.48	-29.1975	1287.61	-94.9541	-3494.71
15	0	4	45	34253	27562	2163.08	1073.91	65.9153	-539.339	-3280.38
16	1	4	45	34253	27562	4953.02	-176.304	-1005.24	1442.65	-4733.21
17	0	2	30	54000	45756	2875.09	1326.25	2663.34	-722.961	-4427.34
18	1	2	30	54000	45756	4875.41	206.114	2353.4	902.268	-5414.73
19	0	4	30	54000	45756	3498.53	1462.23	923.182	330.013	-5183.9
20	1	4	30	54000	45756	6206.03	117.763	59.8199	2720.06	-6590.35
21	0	2	45	54000	45756	2908.09	1272.16	2408.67	-385.728	-4699.65
22	1	2	45	54000	45756	5085.67	49.5932	2010.48	1478.44	-5827.49
23	0	4	45	54000	45756	3587.29	1407.93	498.103	826.505	-5561.05
24	1	4	45	54000	45756	6617.33	-87.8594	-578.37	3668.71	-7211.55

Based on the structural estimates presented in the paper, we simulate willingness to pay for the 5 attributes for 24 people. Income is based on the 25^{th}, median and 75^{th} percentiles of the empirical distribution. Holding all 5 attributes at their sample means, we estimate willingness to pay for a doubling of all attributes but home ownership which is increased from 0 to 100%. Income - price = consumption of non-housing.

Table Nine

Second Stage Product Fixed Effects OLS Regressions

	specification #1		specification #2	
	Coef.	s.e	Coef.	s.e
PUMA per-capita murder rate	-5.096	3.301	-3.365	4.360
miles to the cityhall	-0.083	0.098	-0.039	0.163
% of PUMA who are owners	1.550	3.011	-1.322	4.186
constant	1.281	1.770	1.890	1.857
observations	87		88	
R2	.37		.07	

8 product fixed effects included, standard errors are corrected for PUMA clustering.

Center City Sample Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Product fixed effect	88	.0399042	3.681	-25.36481	7.323601
murders per 1,000 people	88	.2691897	.17124	.0354998	.6021985
miles from city hall	88	5.224006	3.146516	.9998525	12.55009
% of PUMA who own	88	.5600026	.1279661	.2770061	.7205254

Table Ten

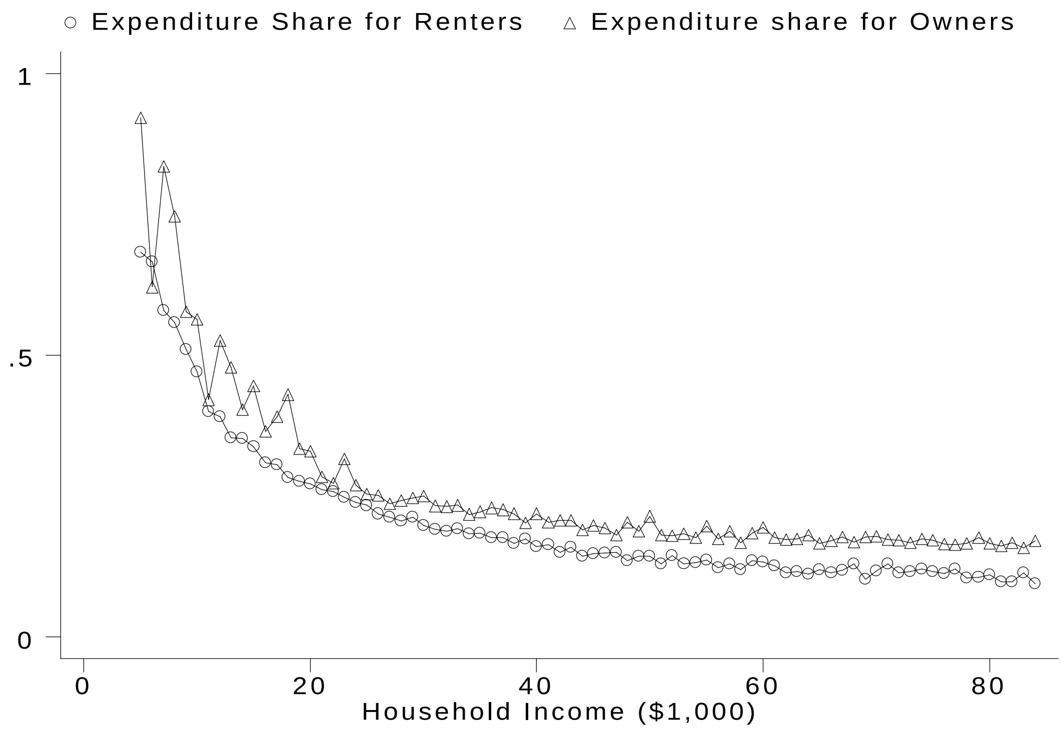
Philadelphia Migrant Willingness To Pay by Demographic Group

Demographic Group						Willingness to Pay in $ per year for an increase of;	
type	white	pers	age	income	cons	to avoid a .0003 probability of murder	if ownership in the community increases by 50 percentage points
1	0	2	30	18700	13200	-1395.91	769.51
2	1	2	30	18700	13200	-3130.8	1941.84
3	0	4	30	18700	13200	-1421.65	784.98
4	1	4	30	18700	13200	-4380	2993.5
5	0	2	45	18700	13200	-1462.35	809.54
6	1	2	45	18700	13200	-3561.23	2280.92
7	0	4	45	18700	13200	-1524.32	847.18
8	1	4	45	18700	13200	-5667.02	4341.49
9	0	2	30	34253	27562.5	-2924.8	1612.82
10	1	2	30	34253	27562.5	-5045.71	2975.01
11	0	4	30	34253	27562.5	-2958.97	1633.35
12	1	4	30	34253	27562.5	-5762.46	3479.93
13	0	2	45	34253	27562.5	-3014.8	1667
14	1	2	45	34253	27562.5	-5380.09	3207.51
15	0	4	45	34253	27562.5	-3075.88	1703.93
16	1	4	45	34253	27562.5	-6317.62	3888.53
17	0	2	30	54000	45756	-4862.25	2681.52
18	1	2	30	54000	45756	-7546.19	4377.42
19	0	4	30	54000	45756	-4906.85	2708.32
20	1	4	30	54000	45756	-8203.5	4820.36
21	0	2	45	54000	45756	-4980.45	2752.65
22	1	2	45	54000	45756	-7895.42	4611.32
23	0	4	45	54000	45756	-5051.44	2795.52
24	1	4	45	54000	45756	-8701.25	5163.64

Using the results from Table Nine and the structural estimates presented in the paper, we simulate willingness to pay for murder and PUMA ownership for 24 people. Income is based on the 25^{th}, median and 75^{th} percentiles of the empirical distribution. All other product attributes are held at their sample means.

Figure One

Housing Expenditure Shares



Philadelphia Migrants Expenditure Shares for Housing