NBER WORKING PAPER SERIES

A DIRECT TEST OF THE "EXPLANATION" FOR INCOMPLETE STRATIFICATION IN VERTICAL SORTING MODELS

Jacob Fishman
V. Kerry Smith

Working Paper 19387
http://www.nber.org/papers/w19387

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2013

This paper was previously circulated as "Heterogeneous Tastes and Incomplete Stratification in Sorting Models: What do Surveys Tell Us?" This material is based upon research supported by the National Science Foundation under grantsBCS-1026865 and DEB-0423704 CAP LTER. Thanks are due to Shauna Mortensen for assistance in preparing this manuscript, to Sharon Harlan, Michael McLean, Marcia Nation, and Carlos Valcarcel for assistance with the PASS survey data, and to Michael Kaminsky and Carlos Silva for assistance with the test score data for our education quality measure and air pollution measure respectively. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

A Direct Test of the "Explanation" for Incomplete Stratification in Vertical Sorting Models
Jacob Fishman and V. Kerry Smith
NBER Working Paper No. 19387
August 2013, Revised March 2014
JEL No. D58,H4,Q51

ABSTRACT

This paper uses the 2011 Phoenix Area Social Survey to evaluate the plausibility of the assumptions made with pure characteristics or vertical sorting models to rationalize incomplete stratification of households across local communities by income. The analysis with a well-recognized index of environmental attitudes, the New Ecological Paradigm (NEP), confirms the correlations in equilibrium outcomes implied by these models. As a result, it supports the role of differences in the tastes for public goods as an explanation for the sorting outcomes.

Jacob Fishman
Department of Economics
Arizona State University
Tempe, Arizona 85287
fishman.jk@gmail.com

V. Kerry Smith
Department of Economics
W.P. Carey School of Business
P.O. Box 879801
Arizona State University
Tempe, AZ 85287-9801
and NBER
kerry.smith@asu.edu

1. Introduction

Recognition that individuals sort among metro areas and between the neighborhoods within a city has had a profound effect on both the structural and the reduced form models in applied micro-economic applications. The structural applications combine market and nonmarket outcomes of these equilibrium adjustments, together with a set of maintained assumptions, to estimate consumer preferences. The most active areas of this research have been in public and environmental economics. In these applications, data from housing markets are used to estimate preferences for local public goods.[1] An important class of these models, the pure characteristics framework, yields two specific predictions for the characteristics of the equilibrium in the presence of sorting. Prices defined for "synthetically equivalent" homes across the discrete number of communities in these households' choice set should, in equilibrium, be ranked in the same way as an index of the different amounts of local public goods provided in these communities. Second, depending on what is assumed about unobserved tastes for these local public goods, the model predicts people with different incomes will stratify by income across communities. The highest income households will be in communities providing the most public goods. To the extent people have *different* preferences for the local public goods, the stratification will not be perfect.

This qualification to the second outcome, namely incomplete stratification, is an important aspect of generalizations to the sorting framework developed by Epple and Platt [1988]. The empirical record has found incomplete stratification by income. All applications of the pure characteristics model have supported the condition required by their model to "produce" incomplete stratification—a negative correlation between household income and the latent taste

[1]See Kuminoff et al. [2013] for a review of the sorting literature and Parmeter and Pope [2013] for a comparable overview of the quasi-experimental method along with discussion of some applications relevant to urban, public and environmental policies.

parameter for local public goods. This parameter is important because it is assumed to capture preference heterogeneity. Consistency of the estimates with this requirement should not be surprising. The models are estimated using moments of the observed distributions of income across communities together with measures for housing prices and the public goods index. The observed income distributions reflect incomplete stratification. If the price index is serving the role assumed, as an aggregate index for the local public goods in equilibrium, then the only way the estimated model can reconcile the maintained structure of preferences with these data is to produce a negative correlation.

This paper proposes a *direct test* of this explanation for incomplete stratification. That is, *we do not* restrict the relationships used to estimate the correlation between income and a taste measure so that they *must* reconcile diverse measures of the income distributions across the neighborhoods assumed to comprise the locational equilibrium. Instead, we develop a separate estimate of the taste index for local public goods from the psychology literature on environmental attitudes. Our index, the New Ecological Paradigm (NEP), is the most widely accepted index of environmental attitudes in this literature. It was developed over 35 years ago and, as Hawcroft and Milford [2010] observe, "The universal nature of the beliefs measured by the NEP scale may explain why it has become the most widely used measure of EA (environmental attitudes) since its publication in 1978." (p. 144, with the parenthetical phrase added).

Our approach exploits the ascending bundles property of the pure characteristics model but does not impose it on our testing framework. We also restrict the data used in our test to consider outcomes at the community level because this aggregation level is the one that is relevant to the negative correlation which reconciles incomplete stratification with the pure

characterization model. The sample size for our test is the number of neighborhoods assumed to comprise the choice set. In our case it is under 40 so our approach further reduces the chances we would find support for a negative correlation.

Our findings are dramatic. We would conclude, based on a simple comparison that is confined to the estimated mean NEP and mean household income at the neighborhood level, that our taste index and income are unrelated. Once we control for community level public goods, however, income and the taste index are significantly negatively correlated. The control used to account for local public goods is the preferred price index developed by Sieg et al. [2002] and estimated using micro level housing sales data for the neighborhoods included in our analysis. We also find it is consistent with ascending bundles condition using a simple comparison. Thus, we have exceptionally strong confirmation for the two primary insights of the pure characteristics version of the sorting model but did *not* maintain either the strong distributional assumption or the preference specification associated with past applications of the structural model.

Section two provides some historical context describing why sorting logic has become so important in the current literature, as an empirical realization of Tiebout's [1956] logic. It also reviews the basic elements in the pure characteristics model. Section three describes the three interrelated steps that were necessary to develop our unique test of the sorting logic. Section four discusses the results and their implications as a "validation" of the sorting logic. The last section discusses the more general implications of our research design and findings.

2. Pure Characteristics Sorting Model

A. Context

Sixty year ago Samuelson [1954] published his classic paper on the pure theory of public expenditures, arguing that no "market type" solution could determine the economically efficient level of expenditures for pure public goods. Two years later, Tiebout [1956] argued that, for local public goods, spatial mobility provided the counterpart of the market trips for private goods. For local public goods he observed:

> "Just as the consumer can be visualized as walking to a private market place to buy his goods, prices of which are set, we place him in the position of walking to a community where the prices (taxes) of community services are set. . . . There is no way in which the consumer can avoid revealing his preferences in a spatial economy." (P. 422 emphasis added)

Both papers had a transformative influence on the ensuing six decades of research in public economics .However, judging by citations, Tiebout's framework has been more influential, with nearly a two to one advantage in terms of citations.[2]

In the current literature sorting models are the empirical embodiment of Tiebout's contention that consumers reveal their preferences for local public goods through their locational choices in a spatial delimited economy. Dennis Epple and his collaborators (i.e. Epple and Platt [1998], Epple and Romano [1996], and Epple and Sieg [1999]) are responsible for many of the contributions that helped to explain how these empirical insights could be derived from households' decisions to locate among a finite set of communities. They demonstrated that there were predictable implications for the distribution of income across communities. Moreover, there were also predictions for the ranking of prices for a

[2] Google Scholar accessed 3/12/14 records 11,869 citations to Tiebout and 6,132 to Samuelson.

standardized unit of housing across communities in relation to an index for the components of the local public goods in those communities.

B. Overview of Basic Model

To appreciate how this logic works consider a simple model. Assume that in a given area, there are a finite set of neighborhoods. In empirical studies these neighborhoods have been defined as school districts or census tracts. Each house is assumed to be represented as a set of characteristics, such as interior square feet, bath rooms, lot size and so forth. Local public goods and amenities are conveyed to those living in the home by virtue of the home's location in one of these neighborhoods. Equation (1) specifies a household's (i) objective function with utility determined by consumption of housing services, h_k, at a location k; a composite numeraire, private good, Z_i;an index of the local public goods and amenities, g_j that varies with each neighborhood, j; the unobservable taste index, α_i which is our focus here; and demographic features of the household, d_i.

The budget constraint, also in (1), has income, m_i, the price of housing $P_{k\in j}$ (i.e. P with the subscript $k \in j$ means the annualized price for house k in community$j$) and the price of the numeraire good normalized to unity.

(1) $$\text{Max } U(\mathrm{Z}_i, h_k, g_j; \alpha_{i,} d_i) \qquad k \in j, \mathrm{Z}$$

$$m_i = \mathrm{Z}_i + P_{k\in j}$$

Households have full information and there are no moving costs. The solution to this choice problem is usually described with an indirect utility function. A specific form for this function is

the usual starting part for most empirical sorting models. The maintained properties of the function are what distinguish the pure characteristics and random utility versions of the model.[3]

Two features of the pure characteristics model are important to our objective. First, all households are assumed to evaluate local public goods and amenities in the same way. The unobserved taste parameter provides the "explanation" for why those with the same income do not select the same location. The second feature is the single crossing condition. This assumption implies the slope of an indifference curve defined in terms of the arguments of the indirect utility function, designated here with the function $v(.)$, in (g, P) space increases monotonically with both income (m) and with the unobserved taste parameter, α, as in equation (2).

$$(2) \qquad \frac{\partial}{\partial m}\left(\frac{dP}{dg}\Big|_{v=\bar{v}}\right) > 0 \quad \frac{\partial}{\partial \alpha}\left(\frac{dP}{dg}\Big|_{v=\bar{v}}\right) > 0$$

This property implies that the equilibrium distribution of households will have specific features for any two neighborhoods. Conditional on stratification by income (for given tastes) the equilibrium implies that if we order the neighborhoods by the equilibrium price of each location, then the index of local public goods and amenities will have the same ordering. This outcome is referred to as the ascending bundles condition and is defined formally in equation (3).

$$(3) \qquad m_{j+1}(\alpha) > m_j(\alpha) \Longrightarrow P_{j+1} > P_j \text{ and } g_{j+1} > g_j$$

The second part of the equilibrium is stratification by income given α and stratification by α given income. Incomplete stratification implies we should observe negative correlations between measures of household preferences for public goods and income, after controlling for all of the local public goods. The price index and index of public goods are linked by the equilibrium. So in equilibrium the price index provides the ideal aggregation for all public goods considered by households.

[3] See Kuminoff, Smith and Timmins [2013] for a more complete discussion of the features of the two modeling strategies.

To date, applications of the model develop estimates for the price index for a "standardized" unit of housing based on transaction data and then use moments of the income distribution or of the housing expenditure distribution (or both) for each community together with that price index and measures for local public goods. The models generally adopt some variation of a constant elasticity specification for the indirect utility function.[4] In this setting the maintained assumption associated with the preference distribution together with normality for the joint distribution of log income and the latent taste parameter for public goods play an important role in "reconciling" the observed income distributions with those predicted by the model. Negative correlation between income and tastes has been the consensus outcome in all applications.[5]

3. Developing our Test

Our analysis focuses on the plausibility of the specification for preference heterogeneity that assures the pure characteristics model will be consistent with incomplete stratification by income. There are three elements that must be assembled to implement our direct test of the logic used in the sorting framework. The first step is to define a set of spatially delineated neighborhoods. Our selection for these neighborhoods was facilitated by the extensive research undertaken as part of the Phoenix Area Social Survey (PASS). This survey was conducted as one of the activities of the NSF sponsored Central Arizona Project (CAP) LTER housed at Arizona State University. The first social survey was conducted in 2006. Our analysis focuses primarily on the second survey conducted in 2011. The 2011 replication started with the neighborhood definitions established in 2006 and added five new neighborhoods.

[4] See Sieg et al. [1004] or Walsh [2007] as examples.
[5] Sieg et.al. [2004], Walsh [2007] and Klaiber and Smith [2012] all found negative correlations in their estimates of pure characteristics sorting models. They were -0.29 to -0.19 for the first study, -0.02 for Walsh, and -0.28 for Klaiber and Smith.

The logic for construction of these areas was based on two criteria: the monitoring of local ecosystems in the Phoenix metropolitan area and the identification of local communities based on demographic criteria including income, ethnicity, and retirement status. The first criterion uses the 204 ecological monitoring sites maintained as part of the CAP-LTER. These monitors study vegetation, soil, and other ecological variables on 30 x 30 meter sample plots distributed over all types of land uses in the study area (see Grimm and Redman [2004]). Initial definitions for neighborhoods were selected after examining aerial photographs of the areas surrounding 101 of the monitoring sites.[6] The remaining 94 sites (101 in residential areas less the 7 eliminated sites, see note #4) were aligned with Census Block groups to identify the socio-economic groups for developing the sampling units based one the second selection criterion. Eight groups were specified, including: low income Phoenix core; low income suburban; middle to high income Phoenix core; middle income suburban; low to middle income fringe areas; high income suburban; high income fringe; and retirement communities. A total of 40 neighborhoods were selected. Five neighborhoods were selected from each group to reflect the demographic composition, the mix of owners and renters, as well as, to match the monitoring data. In 2011, five new neighborhoods were added using the same basic structure, recognizing the areas of population growth in Phoenix. These neighborhoods provide the observational unit for our test. The sorting model predicts distribution of outcomes at the neighborhood level. Thus, we summarize our measures of income and attitudes at this level.

The second step is the development of a price index for a homogenous unit of housing. Here we follow Sieg et al. [2002] and use a hedonic model with housing sales to estimate, as fixed effects, price indexes for the 45 CAP-LTER spatially defined neighborhoods. The last step

[6]Seven sites of the sixteen visited were eliminated because the residents were not close to the plot used for monitoring.

requires an independent measure of tastes for local public goods. In this case we use the index labeled New Ecological Paradigm (NEP). Originally proposed by Dunlap and Van Liere [1978], the NEP was expanded and updated (Dunlap et al. [1992]) to update the questions used to elicit five dimensions of environmental attitudes.[7] They are: beliefs in limits to growth; sentiments against a focus on a human centered view of the environment; concern about the fragility of nature's balance; attitudes that question a view that implies humans are exempt from natural constraints; and concerns about the likelihood of an ecological catastrophe.

There are several reasons for selecting this index. First, as documented by Hawcroft and Milfont [2010] it has been widely used both in the U.S. and in international surveys and is regarded as a reliable indicator of environmental attitudes. Second, it has also been included in both stated and revealed preference studies of the economic tradeoffs people would make to enhance specific dimensions of environmental resources. In these applications it has been an effective indicator of preference heterogeneity (see Kotchen and Reiling [2000] and Kotchen and Moore [2007] as examples.) As we discuss below, it is also correlated with attitudes associated with other local public services at the respondent level. Finally, as a practical matter, confidentiality concerns generally limit the ability of researchers to attach survey responses to small neighborhoods. The second author's role in the design of the PASS 2011 survey allowed access in a way that assured confidentiality but allowed the spatial identification of the responses at the PASS neighborhood level. This last feature is essential to being able to construct our *independent* test.

Our test requires the spatially explicit delineation of the attitude and income measures to neighborhoods at a scale consistent with the housing sales data. The independently estimated

[7] The original NEP scale had 12 items (8 pro-trait and 4 con-trait) and was based on a four point Likert scale using test identifiers of strongly agree and strongly disagree for the anchors. The new items expanded the scope based on comments and is based on 15 questions with a 5 point Likert scale.

price index captures the effects of local public goods and amenities associated with these neighborhoods, and our summary measures for the NEP and income provide the information at the same neighborhood level.

A. PASS Survey

The PASS survey was administered by the Institute for Social Science Research at Arizona State University from May 26, 2011 to January 6, 2012. The target population was heads of households aged 18 or older who lived in one of the 45 neighborhoods defined based on the ecological and socio-economic criteria that underlie the sampling design. 806 completed surveys for a minimum response rate of 43.4 percent.[8]Figure 1 overlays the locations of the neighborhoods with the Census Blocks for the Phoenix metropolitan area.

B. Price Index for Housing

Table 2 provides the estimated hedonic equation used to develop price indexes for a standardized house in each neighborhood as well as descriptive statistics for the sample. We used housing sales prices for single family homes sold from 1995 to 2008 in 39 of the 45 PASS neighborhoods. The omitted six did not have sufficient housing sales. All of the home sales take place before the second round of the PASS survey was undertaken. A semi-log function (in the housing price), specified to include square feet of living space, number of stories, bathrooms, age of the home, presence of garage, pool, number of rooms, and lot size, along with sale year fixed effects and PASS neighborhood fixed effects, was estimated with 20,373 observations for the sales transactions in the PASS neighborhoods.[9] Given the maintained assumptions of the pure

[8]The minimum response rate is the number of complete interviews divided by the number of interviews (complete plus partial) plus the number of non-interviews (refusal and break-off plus non-contacts plus others) plus all cases of unknown eligibility.

[9]We also investigated the effects of reducing the sample sale to 2007, due to the housing downturn in Phoenix and this did not affect our conclusion.

characteristics model, a price index based on these neighborhood fixed effects will, in equilibrium, reflect all the local public services conveyed by each neighborhood's location.

This argument relies on the ascending bundles condition. To assess the consistency of prices and local public goods with this condition we assembled a few measures of local public goods that would be relevant at the spatial scale of the PASS neighborhoods. These include: a measure of local air quality, a measure of the quality of local public schools, and a fixed effect indicating whether the PASS neighborhood is mesic or xeric landscape.[10] To develop a composite index of these three services we used the first factor from a factor analysis of these variables. Figure 2 compares the rank of the price index with that for the first factor used as an index for these local public services. Using a 95% confidence interval for the predictions from a simple regression of the ranks for price on the ranks for this factor reflecting local public goods as a gauge of consistency we find that all but one of the pairs of ranks for the PASS

[10]The education is the arithmetic mean of test scores in reading and math for 2003-2007 for grades 7, 8 and 9. The scores are matched using the property IDs for each house in the school district with the scores. The houses are then matched to the PASS neighborhoods and the average reflects a weighted average of the houses associated with each district that are in a PASS neighborhood. The process used for the air quality measures was somewhat different because the number of monitors was more limited and the records were averages of the daily readings from December 8, 2010 to December 8, 2011 for the PASS neighborhoods that could be matched with the closest monitor. Maricopa County contains 23 air monitoring stations. During this period some stations were closed or not provided by the online data. We considered both particulates (PM10) and the AQI. Monitoring stations identified as: Buckeye, Central Phoenix, Durango, Dysart, Glendale, Greenwood, Mesa, North Phoenix, South Phoenix, South Scottsdale, West 43rd, West Chandler, West Phoenix, Zuni Hills were considered in establishing our matching to PASS neighborhoods. PM10 was used for our factor analysis index. Information on landscape characteristics is obtained from remote sensing data used by Stefanov et al [2001] which classified satellite imagery in the Phoenix area into 12 unique categories. Their classification system analyzed differences in reflectivity to assign one of 12 land cover types to 30x30 meter squares covering our study area. The land cover types include cultivated vegetation, cultivated grass, vegetation, fluvial and lacustrine sediments (canals), water, undisturbed, disturbed soil with agricultural water rights, compacted soil, disturbed (commercial/industrial), disturbed (asphalt and concrete), disturbed (mesic residential), and disturbed (xeric residential). From these categories we create a measure for xeric and irrigated landscape by aggregating the individual categories. By comparing the land cover classification to aerial photography of residential lots, we choose to combine the categories cultivated vegetation, water, and mesic residential to form an irrigated classification and to use the xeric residential category as a "dry" classification. Our classifications are overlaid on parcel level GIS maps assigning a classification as irrigated, xeric, or other to each parcel based on an intersection of parcel centroids. A neighborhood measure of green landscaping is created by aggregating the house specific landscape measures to form subdivision wide indicators of predominately irrigated and predominately xeric subdivisions. For a subdivision to be characterized as mesic, at least 75% of the homes in the subdivision must be identified as irrigated. For a xeric characterization, at least 90% of the homes must be labeled as dry.

neighborhoods fall within the confidence interval. This outcome for readily observable local public goods suggests broad agreement with the ascending bundles condition. Thus, we will assume the price index offers the best summary of all observable and unobservable (to the analyst) local public services.

C. NEP Measures for PASS

The new version of the NEP index is derived from 15 Likert scale (five points) questions. Response categories for each item are "strongly agree", "somewhat agree", "unsure", "somewhat disagree", and "strongly disagree." As Clark et al. [2003] explain, before combining the items into a single index it is desirable to check for the internal consistency of the responses. We follow their strategy in developing this assessment and use variations on their three indicators: the simple correlation between each item's response and overall NEP index, Cronbach's alpha coefficient to gauge the level of reliability of a single scale summarizing the expressed attitudes, and, finally, using the item responses to develop a factor analysis, we consider the size of the first factor loading and its contribution to the variance associated with this factor.

Table 1 summarizes the 15 questions in the NEP coded so the 1 to 5 scores are consistent with the way ratings contribute to the NEP. This format implies the questions where "strongly disagree" would be consistent with a high attitude level for environmental objectives were coded as 5. Similarly, questions where "strongly agree" would be consistent with a high attitude level for environmental objectives were coded as 5. That table reports the percentage of respondents providing the score 1 to 5 for each item and the number of respondents answering the question. Measures for the other three gauges of consistency are also reported in the table. The simple correlations with NEP range from .32 to .66. Cronbach's alpha is measured as the square of the correlation between the measured scale (the sum of the item scores) and the first factor from a

factor analysis. Our estimate for alpha is .796 which is larger than what Clark et al. report and consistent with using NEP at the neighborhood level as a gauge of environmental attitudes. The factor loadings for each question in the NEP index for the first factor range from .20 to .62 across the items. The first factor accounts for 80 percent of the variance among items, suggesting a single index of attitudes offers a reasonable summary of these attitudes.

As we noted at the outset, the NEP is widely accepted as a measure of environmental attitudes. We are proposing to use it here as an indicator of an individual's preferences for local public goods. Our hypothesis is that it provides a measure that should be correlated with the latent variable that underlies the pure characteristics model's explanation for observing incomplete stratification of households by income across communities ranked by the public goods they provide.

The primary objective of the PASS survey is to gauge environmental attitudes and beliefs. Nonetheless, there are other issues raised. To gauge the plausibility of our use of the NEP as indicative of preferences for local public services we selected two of the attitudinal variables collected in the PASS survey that are not included in the NEP. These questions involve the quality of local public schools and crime. The remaining questions ask about other issues such as public transportation but describe the issue as a strategy to reduce pollution. As a result they are not separate for environmental services. The text for the questions we used is given in Table 3. Our empirical analysis recodes the variables into fixed effects. In the case of local schools, the effect is coded as 1 if a respondent is very dissatisfied with local education and zero all other responses (aside from the don't know and non-response that are coded as missing). In the case of crime, a coding of 1 corresponds to responses labeling crime as a big problem. Zero corresponds to responses of a little problem or no problem.

Clearly, these questions do not provide ideal indicators of attitudes toward other local public services because they ask for an assessment of an individual's concern about the issues for the neighborhood where they live. As a result, one might argue they involve a composite of two things: a judgment about local conditions and an assessment of how important the services are. While it is reasonable to assume the two considerations are correlated with individual preferences for services related to both local conditions and the level of concern about the type of services, we would have preferred a pure attitude question. Unfortunately, using other surveys with more carefully directed attitude measures do not have the spatial resolution that is a key element in our test.

Table 4 reports four regression equations using the individual survey responses, rather than average outcomes used for our test in the next section. Here we report results for two regression models using the full NEP in columns (1) and (2) and two using Kotchen and Moore's [2007] reduced NEP index (labeled NEP1 based on five questions that are identified in Table 1 with an asterisk alongside the questions). These findings are in columns (3) and (4).

Columns (1) and (2) regress each NEP measure on household income. Columns (2) and (4) introduce the other two indexes for attitudes toward public services into these basic models. Consistent with other attitudinal research, individuals who score highly on environmental attitudes are likely to have concerns about other public services. Often they are positively correlated but this outcome depends on the service. Our results with the full NEP and the reduced NEP are consistent with this general outcome. It is also important to note the link between income and either measure of NEP at the micro level is not improved by accounting for these variables or other demographics. These findings are consistent with the conclusion that NEP is not serving as a proxy for household income or status. Overall, our findings suggest that

both NEP measures reflect some of the heterogeneity in consumer preferences for a wider set of public services than those exclusively associated with environmental amenities.

4. A Direct Measure of Association Between NEP and Income

The hypothesis that motivated our analysis of housing sales and of PASS respondents' attitudes at the micro level is fundamentally about the properties of *the distributions* in each PASS neighborhood that characterize residents' tastes and income. Thus, neighborhood level summary statistics provide the basis for our assessment of whether average household income and an average taste index are negatively correlated across neighborhoods. To assure the comparison controls for differences in the level of local public goods across neighborhoods we also need some means for conditioning for the levels of these services as well. Our neighborhood level price index provides this control. Thus, while the data requirements to conduct our independent test of the pure characteristics model are quite demanding, the test is straightforward.[11] We simply regress the mean NEP on the mean household income and the neighborhood price index. Table 5 provides the results.

Column (1) uses only average household income at the neighborhood level as a determinant of the average NEP. Here our findings parallel what we found with the micro data. In column (2) we add the neighborhood price index as a control for variations in local public services across neighborhoods. The change in our conclusions is striking. Mean household income is now negatively related to NEP and significant at the five percent level. The price index is also a significant influence to NEP implying that this index of public goods is influential to the

[11] The key requirement is the consistent spatial resolution for measures of attitudes, household income and prices as the aggregator reflecting local public goods. A partial regression coefficient between any two variables measures the correlation between those two variables after the linear effect of other variables have been removed. Our test should focus on the partial correlation between average NEP and average income after the effects of local public goods have been removed. By the Frisch-Waugh-Lovell theorem the coefficient on income in a regression model for NEP that includes the price along with income provides this information.

equilibrium sorting of households, as the theory would maintain. The R^2 for the model increases dramatically as well from .026 to .334 with the addition of one variable to control for differences in local public services. Column 3 repeats the exercise with Kotchen and Moore's reduced NEP (labeled NEP1). In terms of the signs of the effects we see consistency with the results for the full NEP. However the estimates are not statistically significant and the model fit is not much better than when the full NEP is used with average income alone.

In column 4 we consider the persistence of the effects of sorting. More specifically, it is possible to construct a version of the NEP using four questions from the 2006 version of PASS for 34 of the 45 neighborhoods. This is labeled NEP2 (and identified with a "b" alongside the questions in Table 1). We developed this index using the 2011 responses and the 2006 responses and matched the averages by neighborhood. Including this variable along with average income from the PASS 2011 survey and our price index yields this model. The negative partial effect of income, a consistent role for the price index, and the apparent persistence in environmental attitudes are all confirmed. The later finding is also consistent with what we would expect from social interaction models. The other results further confirm the outcomes we would expect with sorting.

Our direct test of the conditions needed for the pure characteristics model to produce incomplete stratification was possible because of a unique feature of the design of the PASS social survey. As we noted earlier, the survey is a part of the activities of an interdisciplinary research program focused on the Phoenix metropolitan area as an urban ecosystem. The Central Arizona Project LTER has a specific mandate to measure features of the desert ecosystem over time. This goal necessarily requires a detailed spatial design for the monitoring network. It was natural, as a consequence, to embed the design of the social survey in this detailed spatial

framework. The extensive data bases for housing sales, with ability to link prices to parcel maps, have transformed hedonic literature. These models permit the estimation of price indexes with a high level of spatial resolution. What is unique about our application is the ability to match the attitude and socio economic data to a neighborhood scale with the same spatial delineation. This level of detail allows us to avoid making the distribution assumptions underlying earlier tests of the negative correlation between taste for public goods and income.

5. Summary and Implications

In this paper we provided the first direct evidence confirming the mechanism for incomplete neighborhood stratification of households by income. No doubt many readers familiar with the structural estimates of sorting models in public, environmental, and urban literatures find the negative correlation between income and unobserved tastes for public goods plausible given the consistent estimates from all applications to date.[12] As a result, there may well be a tendency to ask in response to our finding -- so what? As a result, we answer this question in our closing discussion.

Two implications will be discussed. First, several authors, notably Hendren [2013] and Chetty et al. [2014] have argued that general equilibrium effects of policy require an understanding of how preference heterogeneity influences behavioral responses to each new policy. In Hendren's case the policy is the earned income tax credit and the welfare effects are a composite of gains in well-being to low income earners from the added income (due to the credit) versus the cost of providing that gains (due to the increased tax rates) imposed on the high income earners, given a constant government budget. Both agents' behavioral responses to the policy, recognizing the impact of the policy for the government's budget, contribute to the

[12] See note # 5 for examples.

general equilibrium assessment of welfare gain or loss. In order to evaluate the nature of his *policy elasticity* in advance we need to be able to characterize the preference heterogeneity. It is especially important to describe how this heterogeneity is likely to influence the general equilibrium responses to policy. The negative correlation is an example of what is needed. It was important to the general equilibrium analysis of the distributional effects of Clean Air Act policies described in Smith et. al.'s [2004] use of a sorting model to evaluate the effects of changes in ozone for Southern California.

Chetty, Hendren, Kline, and Saez [2014] provide indirect empirical evidence of the importance, over time, of general equilibrium sorting responses. Based on their results, intergenerational mobility varies substantially across metropolitan areas in the U.S. Three aspects of their results are relevant. Two relate to local public goods and the quality of local schools. Both sets of services are positively related to households' intergenerational mobility. These results would seem to suggest that households who move to metro areas consistent with the predictions of the Tiebout hypothesis would be doing so in part to enhance their children's welfare. However, these authors' evaluation of three migration measures found no link between in-migration, out-migration, or fraction of foreign born in a metro area on intergenerational mobility. All of the migration measures display insignificant correlation with their mobility measures. This would seem at odds with the Tiebout logic and our explanation of how heterogeneous preferences account for incomplete stratification as an outcome in static models. Unfortunately, their simple correlations are subject to the same limitations as our analysis of NEP and income. Without controlling for local public goods, as the mechanism associated with a sorting equilibrium would imply, there are no clear-cut expectations about the relationships we should observe. Thus, our evidence for one metropolitan area using the preferences index and

income across neighborhoods offers a potential indirect explanation for their findings. Thus it plays a role in both Hendren's static model and in a model developed to help to explain the Chetty et al. findings. The message in both papers is that heterogeneity in circumstances and preferences influences the behavioral responses households make to policies intended to enhance well-being. This point may be especially relevant for efforts to improve the well-being of households at the lowest end of the income distribution.

Chetty et al. recognize the opportunity to learn from the spatial differences in the conditions that households face. Their analysis to this point has been based on reduced form statistics. Our research supports a return to the Tiebout logic as part of the challenge they identify at the close of their paper—namely developing models that will help in "understanding why some areas of the U.S. persistently generate higher rates of intergenerational mobility than others . . ." (p. 47) We believe that the differences in the access that different households face to Tiebout's spatial market provides part of their answer.

Second, the support for NEP argues for greater attention to integrating indexes of individual attitudes as new observables for estimating the role of preference heterogeneity in structural models. Such strategies would parallel the efforts of Kotchen and Moore [2007] in the analysis of stated preference surveys. With greater attention to the need for spatial resolution in attitude indexes it should be possible to balance the need to assure respondents of the confidentiality of their answers and provide neighborhood level summary measures.

References

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, (2014). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States" NBER Working paper # 19843, January.

Clark, Christopher F., Matthew J. Kotchen, and Michael R. Moore (2003). "Internal and External Influences on Pro-Environmental Behavior: Participation in a Green Electricity Program." *Journal of Environmental Psychology*, 23: 237-246.

Dunlap, Riley E. and Kent D. Van Liere (1978). "The New Environmental Paradigm." *Journal of Environmental Education,* 9: 10-19.

Dunlap, Riley E, Kent D. Van Liere, Angela G. Mertig and Robert E. Jones (2000). "Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale." *Journal of Social Issues*, 56: 425-442.

Ellickson, Bryan C. (1971). "Jurisdictional Fragmentation and Residential Choice." *American Economic Review*61: 334-339.

Epple, Dennis and Glenn J. Platt (1998). "Equilibrium and Local Redistribution in an Urban Economy When Households Differ in Both Preferences and Incomes," *Journal of Urban Economics* 43: 23-51.

Epple, Dennis and Holger Sieg (1999). "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy* 107(4): 645-81.

Grimm, Nancy B. and Charles L. Redman (2004). "Approaches to the Study of Urban Ecosystems: The Case Study of Central Arizona – Phoenix." *Urban Ecosystems*, 7: 199-213.

Hawcroft, Lucy J. and Taciano L. Milfont (2010). "The Use (and Abuse) of the New Environmental Paradigm Scale over the Last 30 Years: A Meta-Analysis." *Journal of Environmental Psychology,* 30: 143-158.

Hendren, Nathaniel (2013). "The Policy Elasticity" NBER Working paper # 19177, June.

Klaiber, H. Allen and V. Kerry Smith (2012). "General Equilibrium Benefit Analyses for Social Programs." *Journal of Benefit Cost Analysis,* March, 1-50.

Kotchen, Matthew J. and Michael R. Moore (2007). "Private Provision of Environmental Public Goods: Household Participation in Green-Electricity Programs." *Journal of Environmental Economics and Management*, 53: 1-16.

Kotchen, Matthew J. and Stephen D. Reiling (2000). “Environmental Attitudes, Motivations, and Contingent Valuation of Nonuse Values: A Case Study Involving Endangered Species.” *Ecological Economics*, 32: 93-107.

Kuminoff, Nicolai V. 2010. "Recovering Preferences in a Dual-Market Sorting Equilibrium." *Working Paper*.

Kuminoff, Nicolai V., V. Kerry Smith and Christopher Timmins (2013).“The New Economics of Equilibrium Sorting and its Transformational Role for Policy Evaluation.” *Journal of Economic Literature*, 51(4):1-57.

Parameter, Christopher and Jaren C. Pope (2013).“Quasi-Experiments and Hedonic Property Value Methods.” in *Handbook on Experimental Economics and the Environment*, eds. John A. List and Michael K. Price, Northampton, MA: Edward Elgar Publishing.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf and Randy Walsh (2002). “Interjurisdictional Housing Prices in Location Equilibrium.” *Journal of Urban Economics*, 52(1) July: 131-153.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh(2004). "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods." *International Economic Review*, 45(4): 1047-77.

Smith, V. Kerry, Holger Sieg, H. Spencer Banzhaf, and Randy Walsh. (2004). "General Equilibrium Benefits for Environmental Improvements: Projected Ozone Reductions under EPA’s Prospective Analysis for the Los Angeles Air Basin." *Journal of Environmental Economics and Management*, 47(3): 559-84.

Stefanov, William L., Michael S. Ramsey, and Phillip R. Christensen(2001).”Monitoring Urban land Cover Change: An Expert System Approach to Land Cover Classification of semiarid to Arid Urban Centers.” *Remote Sensing of Environment*, 77: 173-185.

Tiebout, Charles M. (1956). “A Pure Theory of Local Expenditures.” *Journal of Political Economy*, 64(5) October: 416-424.

Walsh, Randall L. (2007). "Endogenous Open Space Amenities in a Locational Equilibrium." *Journal of Urban Economics*, 61(2): 319-44.

Figure 1

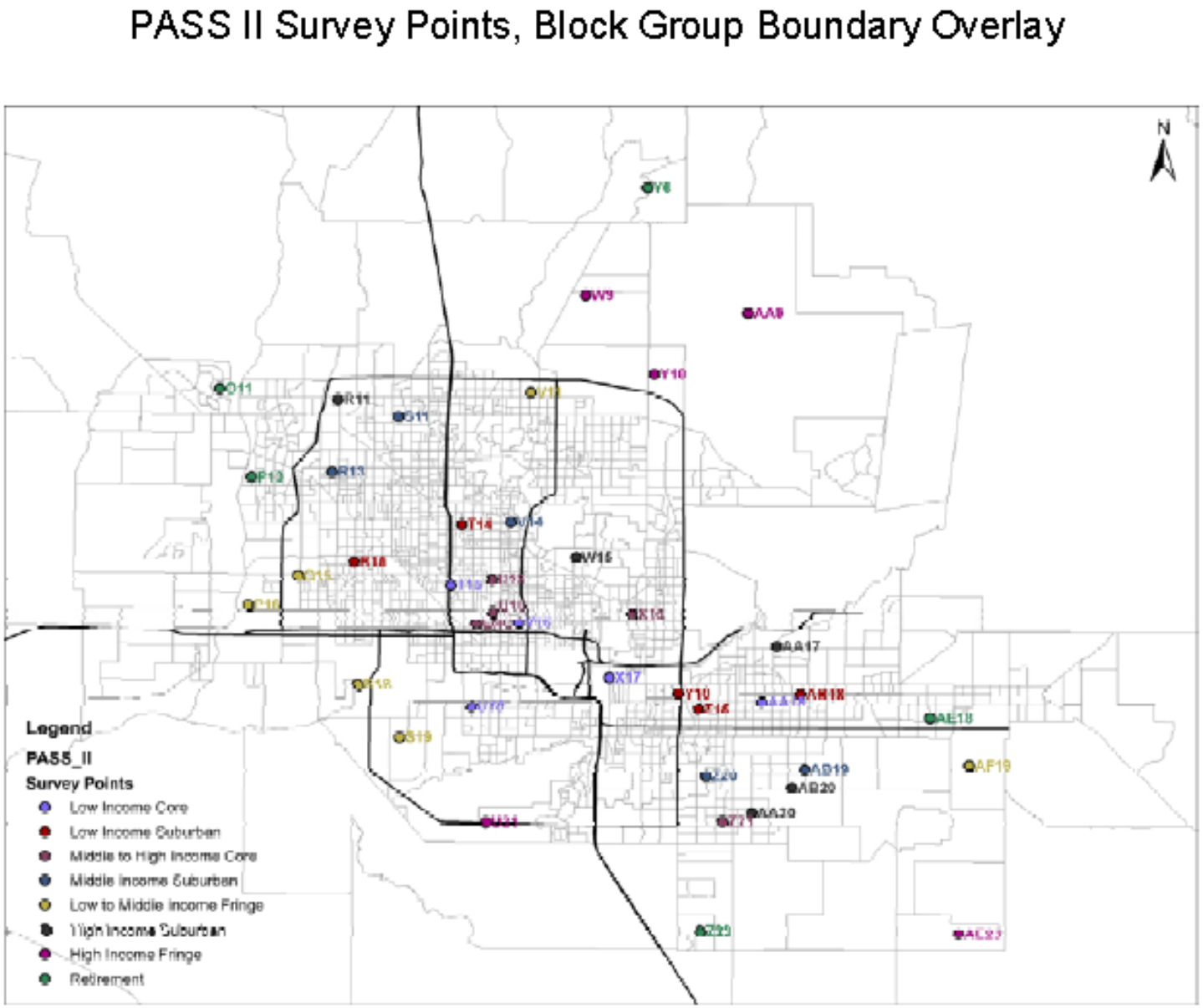
PASS II Survey Points, Block Group Boundary Overlay
N
Legend
PA55_II
Survey Points
Low Income Core
Low Income Suburban
Middle to High Income Core
Middle Income Suburban
Low to Middle Income Fringe
High Income Suburban
High Income Fringe
Retirement

Figure 2: Simple Test of Ascending Bundles for PASS Neighborhoods

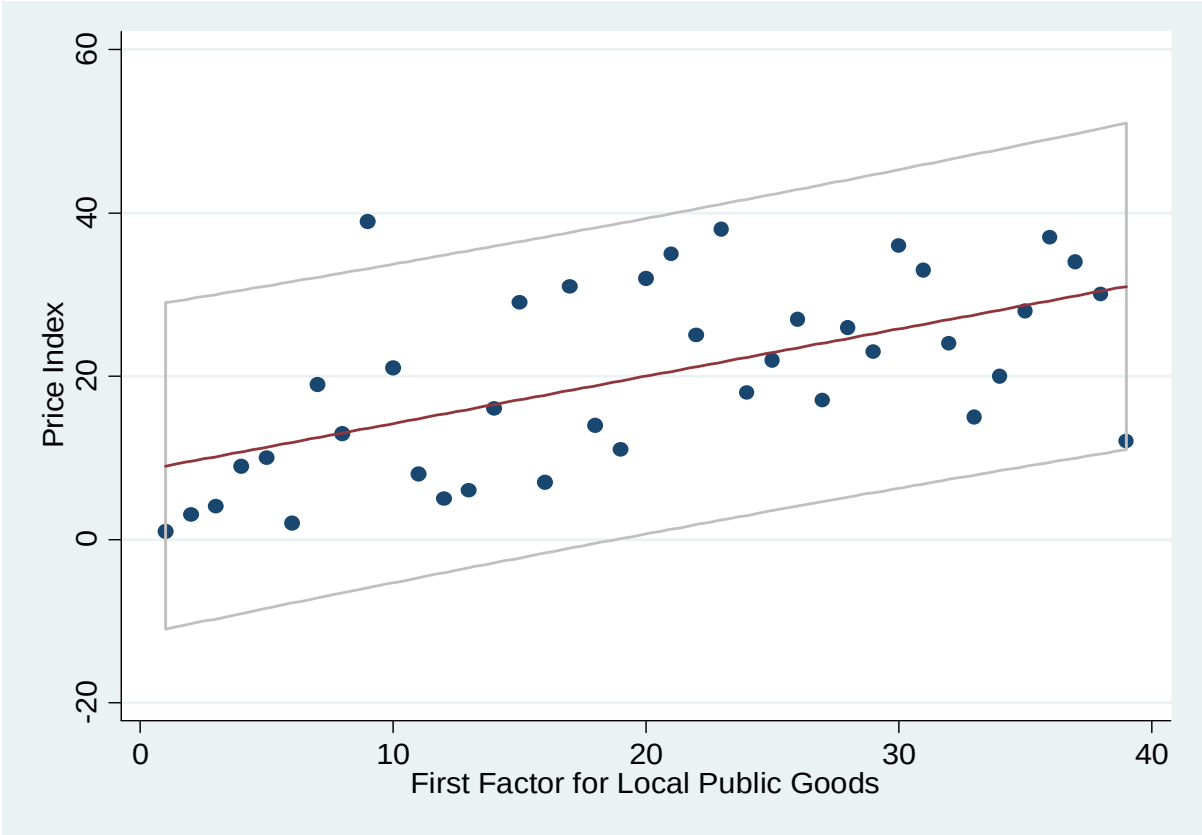

Table 1: Components of the New Environmental Paradigm for the 2011 PASS Survey in Phoenix

		Contribution to NEP							
		1	**2**	**3**	**4**	**5**	**n**	**Loading**	**r_i NEP**
1	We are approaching the limit of the number of people the earth can support.	13.9	21.9	12.6	31.0	20.6	804	0.4643	0.51
2	Humans have the right to modify the natural environment to suit their needs.	9.6	29.5	3.4	32.5	24.9	802	0.3980	0.50
3	b. When humans interface with nature it often produces disastrous consequences.	4.1	15.8	4.5	39.3	36.3	804	0.4907	0.51
4	* Human ingenuity will insure that we do NOT make the earth unlivable.	13.4	34.0	12.0	25.7	15.0	801	0.3304	0.44
5	Humans are severely abusing the environment.	5.8	10.6	2.6	40.4	40.6	805	0.6241	0.61
6	The earth has plenty of natural resources if we just learn how to develop them.	34.2	39.6	4.0	13.4	8.6	804	0.3147	0.42
7	*, b. Plants and animals have as much right as humans to exist.	5.7	9.5	1.7	27.3	55.8	803	0.5002	0.50
8	* The balance of nature is strong enough to cope with the impacts of modern industrial nations.	7.3	25.2	9.3	30.2	27.9	805	0.5633	0.62
9	b. Despite our special abilities and humans are still subject to the laws of nature.	2.2	3.7	3.0	31.8	59.2	802	0.2647	0.32
10	*, b. The so called "ecological crisis" facing humankind has been greatly exaggerated.	14.0	25.5	7.7	26.1	26.6	804	0.6155	0.66
11	* The earth is like a spaceship with very limited room and resources.	14.4	22.5	5.6	36.4	21.1	801	0.4859	0.52
12	Humans were meant to rule over the rest of nature.	13.1	20.3	5.5	25.7	35.4	802	0.4368	0.53
13	The balance of nature is very delicate and easily upset.	4.3	15.3	6.0	37.0	37.4	805	0.5699	0.55
14	Humans will eventually learn enough about how nature works to be able to control it.	9.2	24.2	7.5	28.5	30.6	801	0.1970	0.34
15	If things continue on their present course, we will soon experience a major ecological catastrophe.	9.7	17.0	9.1	36.6	27.6	804	0.6009	0.58

Table 2: Hedonic Price Equation for CAP-LTER Neighborhoods Housing Sales

	Model [a]	Summary Statistics [b]
Living space (hundreds of sq. ft.)	.03	19.88
	(16.27)	(7.43)
Lot size (acres)	.67	0.21
	(17.16)	(0.17)
Stories	-.04	1.25
	(-6.24)	(0.43)
Bathrooms	.05	2.66
	(8.86)	(0.86)
Age of home	-.01	13.35
	(-13.05)	(17.45)
Has garage (0, 1)	.03	0.94
	(3.09)	(0.23)
Has pool (0, 1)	.04	0.31
	(7.78)	(0.46)
Living space ^{2}	$-.01x10^{-2}$	—
	(-7.63)	
Lot size^2	-.09	—
	(-7.63)	
Age^2	$-.01x10^{-2}$	—
	(8.96)	
Sale year fixed effect	yes	24,722 [c]
		(30,907)
Neighborhood fixed effect	yes	—
No. of Observations	20,372	
R^2	0.999	

a. The numbers in parentheses are t-statistics for the null hypothesis of no association. The model is a semi-log in the logarithm of the housing price. It was restricted to exclude an intercept.

b. The numbers in each row are the means for each variable and those in parentheses are the standard deviations.

c. This is the mean for the nominal price of houses in the sample with sales from 1995 to 2008. The value in parentheses is the standard deviation.

Table 3: PASS Attitude Questions [a]

1. Quality of Schools

For each of the following items please indicate whether you are very satisfied, somewhat satisfied, somewhat dissatisfied or very dissatisfied with that item or features In YOUR NEIGHBORHOOD.

1. Very satisfied (desirable)
2. Somewhat satisfied (desirable)
3. Somewhat dissatisfied (Undesirable)
4. Very dissatisfied (undesirable)

98. Don't know

99. Refuse to answer

2. Crime

Some people have mentioned the following problems in their neighborhoods. Please indicate whether each item is a big problem, a little problem, or not a problem at all IN YOUR NEIGHBORHOOD.

1. Big problem
2. Little problem
3. Not a problem

98. Don't know

99. Refuse to answer

a. The order of these two questions was randomized each within a separate, wider set of attitude questions. We selected the one question from each set that was most closely related to an attitude toward local public services.

Table 4: NEP and Attitudes Toward Other Public Services[a]

	NEP		NEP1	
	(1)	(2)	(3)	(4)
Household income (ten thousands)	-0.063	-0.089	-0.017	-0.035
	(-1.05)	(-1.11)	(-0.67)	(-1.00)
Attitude and assessment of local school quality		2.390		0.738
		(2.52)		(1.77)
Attitude and assessment of local crime conditions		3.137		1.076
		(3.40)		(2.66)
Male (=1)		-0.841		-0.146
		(-0.92)		(-0.36)
Race (White =1)		0.334		0.012
		(0.33)		(0.03)
Education (College grad = 1)		1.133		0.738
		(1.08)		(1.61)
Intercept	53.255	49.433	17.307	15.875
	(92.10)	(44.60)	(70.01)	(32.69)
Number of observations	670	464	670	464
R^2	.002	.053	.001	0.033

a.The numbers in parentheses are t-ratios for the null hypothesis of no association.

Table 5: The Relationship of the New Ecological Paradigm and Household Income[a]

	Full NEP		Reduced NEP1	Reduced NEP2
	(1)	(2)	(3)	(4)
Household Income (Thousands)	0.011	-0.038	-0.024	-0.015
	(0.98)	(-2.54)	(-1.20)	(-3.05)
Price Index	———	6.963	3.698	1.740
	———	(4.08)	(1.65)	(3.19)
Parsimonious NEP (2006)	———	———	———	0.141
	———	———	———	(1.51)
Intercept	51.51	54.335	4.648	14.337
	(55.34)	(52.08)	(3.39)	(8.79)
R^2	.026	.334	.071	.382
N	39	39	39	34

a. The numbers in parentheses are t-ratios for the null hypothesis of no association.