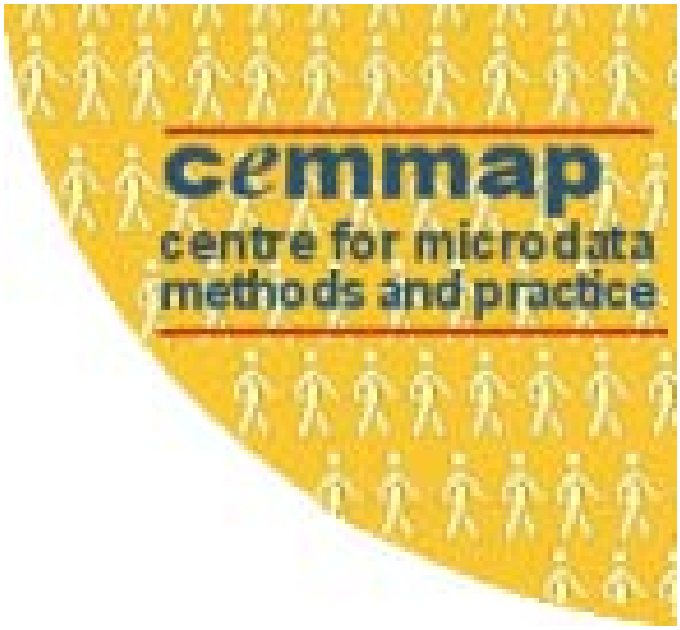

EQUILIBRIUM SORTING OF HETEROGENEOUS CONSUMERS ACROSS LOCATIONS: THEORY AND EMPIRICAL IMPLICATIONS

Lars Nesheim

THE INSTITUTE FOR FISCAL STUDIES
DEPARTMENT OF ECONOMICS, UCL
cemmap working paper CWP08/02

Equilibrium Sorting of Heterogeneous Consumers Across Locations: Theory and Empirical Implications

Lars Nesheim[1]

March, 2002

[1]This paper benefited from discussions with many people. I would like to thank Pierre-Andre Chiappori, Ken Judd, and especially James Heckman. I would also like to thank Renee Adams, Todd Dupont, Ed Glaeser, Lance Lochner, Edward Seiler, Chen Song, George Tolley, and Jean-Pierre Zigrand.

Abstract

This paper develops a model in which a continuum of consumers choose from a continuum of locations indexed by school quality. It computes equilibria that are sustained by a price function that matches consumers to different locations based on their willingness to pay for school quality. In equilibrium each location is inhabited by a set of people with varying levels of education, ability, intensity of preference for education, and income. The distributions of characteristics within each location are determined by the structural elements of the model.

The paper also develops a set of computational algorithms that solve several complex numerical problems. These problems include the calculation of a number of difficult integrals, the calculation of asymptotic approximations to those integrals, the solution of an implicitly defined differential equation that depends on the integrals previously calculated, and the maximization of a likelihood function that depends on the solution of the differential equation.

Finally, this paper demonstrates how the equilibrium implications of a structural economic matching model can be used to solve two important

econometric identification problems. First, it is likely that regressions that seek to estimate the effects of school quality on educational outcomes produce biased and inconsistent estimates because people choose where their children go to school. The model in the paper solves this problem by using a consumer location choice equation and an equilibrium pricing relation to create a valid instrument for the school quality variable. Second, hedonic estimation problems in a single market are unidentified because the marginal price function is unknown or collinear with the level of the product demanded. This paper solves this problem by exploiting the restrictions that equilibrium in the sorting economy imposes on the equilibrium price function. The equilibrium price equation introduces a non-linearity into the system that is sufficient for identification.

1 Introduction

There is large empirical literature that seeks to measure the importance of neighborhood characteristics in the production of individual outcomes. This literature begins with the hypothesis that many outcomes are produced not only by an individual's observed and unobserved background characteristics, but also by the characteristics of the neighborhood in which a person lives. For instance, the average education of the residents of a community along with an individual's own parents' education and the individual's own ability might determine the educational outcome of that individual. Working with similar hypotheses researchers have sought to use regression analysis to explain a large number of social outcomes including educational outcomes, job market outcomes, fertility behavior, and criminal behavior. Important surveys of the results of this research can be found in Brock and Durlauf (2000), Haveman and Wolfe (1995), and Jencks and Mayer (1991).

For the most part, researchers have sidestepped a major issue that affects the interpretation of their empirical results. They have not dealt wholeheartedly with the fact that people choose their location and that as a result it is very likely that the unobservable characteristics that enter the production functions for individual outcomes are correlated with observable neigh-

borhood characteristics. If ability increases the productivity of neighborhood quality in producing education, then people who know they have high ability children will move to high quality neighborhoods. This could result in higher mean unobserved ability in high quality neighborhoods. Alternatively, if parents of low ability children perversely value education very highly, then they will move to high quality neighborhoods, outbidding the parents of high ability kids. This could result in lower mean unobserved ability in high quality neighborhoods. In short, when people choose where to live, there is little that can be said about the distributions of unobservables across locations unless one understands the sorting process.

Researchers have recognized that their empirical neighborhood effects models lack a mechanism describing neighborhood sorting, but they have not been able to specify an economic model that describes the location choices of heterogeneous consumers and develop the implications that such a model has for what types of people live in each location in equilibrium. Nor have they been able to embed such an economic equilibrium framework in an empirical setting.

In this paper, I tackle this problem by developing a model in which a continuum of consumers choose from a continuum of locations indexed by

school quality. I compute equilibria that are sustained by an equilibrium price function that separates consumers into different locations based on their willingness to pay for school quality. I show how sets of consumers are sorted into locations so that in equilibrium, each location has a distribution of people with varying levels of education, ability, intensity of preference for education, and income. The distributions of characteristics within each location are determined by the structural elements of the model.

An important component of the model is a set of computational algorithms that solve several complex numerical problems. These problems include the calculation of a number of difficult integrals, the calculation of asymptotic approximations to those integrals, the solution of an implicitly defined differential equation that depends on the integrals previously calculated, and the maximization of a likelihood function that depends on the solution of the differential equation. The algorithms I develop allow me to analyze the theoretical and empirical properties of the equilibrium sorting model. In addition, they stand on their own as one of the major contributions of this paper.

The model bears a significant resemblance to hedonic equilibrium models developed by Tinbergen (1959), Sattinger (1981), Kniesner and Leeth (1995),

Teulings (1995), Epple and Platt (1998), and Epple and Sieg (1999a,1999b). These models also match consumers to locations and find prices that separate people based on their willingness to pay for locational quality. An important aspect of all of these models except Epple and Platt (1998) and Epple and Sieg (1999a, 1999b) is that consumers' valuations of the locations depend on characteristics of the locations themselves, not on characteristics of the equilibrium sets of people at the locations. In contrast, in this paper as in Epple and Platt and Epple and Sieg, valuations of locations depend on the sets of people at the various locations.[1] This is a fundamental trait of interactions-based models and as such is a fundamental element of the analysis in this paper.

This theoretical analysis produces an equilibrium pricing function and a spatial equilibrium that displays rich patterns of sorting across locations. This equilibrium is valuable for several reasons. First, it yields important insights about the determinants of patterns of neighborhood sorting. Second, it clarifies the econometric issues involved in the identification of neighborhood effects and location choice. Third, it produces restrictions that can be

[1]Epple and Platt and Epple and Sieg develop a model with a finite number of locations. In their papers, locational quality depends on expenditures on a public good. The level of expenditures is determined by the set of residents at each location through a voting mechanism.

exploited to achieve identification.

In addition to the pricing function, the equilibrium produces two empirical equations. The first equation is the education production function that describes the relation between individual educational outcomes, individual background characteristics, and locational quality. As I discussed above, its parameters cannot be consistently estimated because locational quality is chosen by the consumer and hence is correlated with unobserved individual traits. The second equation is the locational choice equation describing each consumer's choice of location as a function of the marginal price of the location and the consumer's individual characteristics. This equation can be used to create an instrument for neighborhood quality in the first equation; however, it must be estimated first. This requires solving a hedonic estimation problem which may not be identified because both the choice variable and the marginal price are endogenous and may by linearly related. Thus, it is clear that solving the neighborhood effects estimation problem entails solving two distinct econometric problems: the endogenous regressor problem and the hedonic estimation problem.

First, consider the hedonic estimation problem. The difficulty lies in estimating the structural parameters that describe demand for locational quality

when that demand depends on a marginal price that varies with actual quality chosen. It is well known that when the price function is unknown and the demand relationships are estimated using price data and an arbitrary functional form to approximate the price function, the structural parameters cannot be reliably estimated using single market data (see Rosen (1974), Brown and Rosen (1981), Bartik (1987), Epple (1987), and Kahn and Lang (1988)). In particular, if the demand equation is assumed to be linear in neighborhood quality and the marginal price is also assumed to be linear, the structural demand parameters cannot be identified from estimation of the neighborhood choice equation. In this case, the empirical equation has two endogenous variables which are perfectly collinear. If the marginal price function is assumed to be non-linear, this aids identification by breaking the collinearity. However, this identification is arbitrary since there is no guide as to what non-linear functional form to use.

Ekeland, Heckman and Nesheim (2001) and this paper develop ways to solve this identification problem. They observe that methods like the one described above, do not use all the information that is available in a hedonic equilibrium model; economic equilibrium imposes restrictions on the hedonic pricing function that generically result in identification of the structural

parameters of the demand relationship. Ekeland, Heckman, and Nesheim show in a non-parametric setting with separable preferences that generically the equilibrium marginal price function is not linear.[2] Exploiting this non-linearity is sufficient for identification. In the parametric model developed in this paper, this only requires the numerical solution of an ordinary differential equation (ODE). This non-linearity is not arbitrary because it is derived from the structure of the equilibrium of the model. The distributions of consumer and supplier types that generate the non-linearity can be estimated from data.

Ekeland, Heckman and Nesheim (2001) develops these ideas in the context of a classical one-to-one matching hedonic equilibrium model. They focus on using equilibrium restrictions to estimate the demand for differentiated commodities. In this paper, I develop these ideas in the context of a one-to-many matching hedonic equilibrium model with neighborhood effects. As in Ekeland, Heckman and Nesheim (2001), I study how equilibrium restrictions on the pricing function can be used to identify the parameters of the hedonic location choice equation. Moreover, I show that this is possible even when

[2]That is, they show that the set of economies in which the equilibrium marginal price function is non-linear is a countable intersection of open dense subsets of the space of feasible economies.

data on prices is not available.

I then extend the analysis to show how equilibrium restrictions not only identify the hedonic demand equation, but simultaneously solve the endogenous regressor problem and identify the parameters of the equation describing the effect of neighborhood characteristics on individual outcomes. This result requires the analysis of a non-linear simultaneous equations system that depends on the solution of a differential equation. I show that the conditions under which this system is not identified are not generic in the sense that small perturbations of the structure of the model (starting from a system that is not identified) result in systems that are identified. In section 5.5, I demonstrate the efficacy of these methods by estimating the structural parameters of a sorting economy using both synthetic and real data.

2 Overview

The outline of the paper is as follows. In section 3, I describe the supply and demand conditions of the model and set out the specific economic environment facing the consumer. The model is a model of locational choice in which consumers pay more to live in locations with higher levels of school quality.

Consumers value locations with higher quality schools because school quality is an input into their child's educational outcome. School quality in each location depends on the set of people living there. I assume that the average education of the parents in a location is the relevant measure of school quality.[3] The specific assumptions made are aimed toward creating the simplest possible model that has the rich sorting equilibrium described in the previous section and that can be used for empirical work.

After setting out the basic assumptions, I describe the equilibrium concept that is used in the model, and discuss some basic facts that are true of equilibria in many general models of this type. The equilibrium is a Nash equilibrium in which consumers treat the school quality in each location and the price function as given, and choose their location accordingly. In equilibrium, the quality in each location is consistent with the quality they expect when they make their location choice. I show that there is always a trivial equilibrium in which every location has the same school quality and the same price of zero. Then I show that if there are "separating" equilibria in which

[3]There is a great deal of controversy about what factors influence school quality (Betts (1995), Card and Krueger (1992), Hanushek (1996), Heckman et al. (1996)). In this paper, I remain agnostic about the results of this literature and note that the results in Table 1 show a positive correlation between average parental education at a child's school and the child's subsequent schooling attainment. I seek to explain whether this correlation indicates a causal relationship and how strong that relationship might be.

consumers separate into locations with different average levels of education, these equilibria can be partitioned into unique classes of equilibria that assign each consumer to a unique quality of location. These "quality" sorting equilibria are the central objects of study throughout the remainder of the paper.

In section 4, I analyze the simplest example of the model developed in section 3 and characterize the properties of its equilibrium. In this example, utility is linear in consumption, and I derive a closed-form expression for the equilibrium price function. I also derive a closed form expression that describes the equilibrium distributions of consumer types within each neighborhood. The simplicity of these results provide a clear illustration of how different factors affect both the equilibrium price function and the equilibrium patterns of sorting that are observed in the economy. These results are easy to understand and easy to compute.

After developing these results, I analyze the empirical equations resulting from the linear utility model in section 4.3. These equations are linear functions of the logarithms of the data. They clearly illustrate the primary econometric issues involved in estimating the model. The first issue is the estimation of the hedonic location choice equation. The hedonic equation

contains the logarithm of the neighborhood quality variable as well as the logarithm of the marginal price variable. Both of these are endogenous and in the linear utility model equilibrium both are linearly related. These problems of endogeneity and collinearity must be addressed to estimate the hedonic equation. The second issue is the endogeneity of the school quality variable in the education production function. An instrument must be found for this endogenous explanatory variable. The location choice equation can produce a valid instrument for this variable if it can produce a predicted school quality that is linearly independent of the other regressors in the education production function. The location choice equation resulting from the linear utility model fails this test.

This conclusion leads me to relax the assumptions imposed on the model. In section 5, I study one relaxation in which utility is an exponential function of consumption. This example has no closed form equilibrium price function but allows for richer patterns of sorting than the simple linear example. The equilibrium price function is the solution to a non-standard ordinary differential equation whose approximate solution must be computed using simple but specialized numerical techniques. In computing these approximations, I develop techniques to solve several difficult numerical problems.

First, the differential equation itself depends on the ratio of two complicated integrals that must be approximated quickly and accurately. Moreover, the approximations must vary smoothly with the input parameters and with the state variable of the differential equation. I develop an algorithm that meets these criteria by carefully analyzing the integrand, making a change of variable, and then using Gauss-Chebyshev integration formulas.[4]

Second, in some regions of the state space, the integrals entering the differential equation are not computable using the integration technique outlined above. In these regions, the values of both integrals cannot be distinguished from zero using a finite precision computer. To compute the solution of the differential equation in these regions, I develop asymptotic expansions that approximate the values of these integrals and use the ratios of these approximations to compute the solution of the differential equation.[4]

Third, the differential equation can have singularity points at some points in the state space. I show that when these singularity points exist, the equilibrium price function is not twice continuously differentiable. I develop an algorithm that tests whether singularity points exist, and computes a piecewise twice continuously differentiable approximation to the equilibrium

[4]A comprehensive text discussing these techniques and their uses in economics is Judd (1998).

price function when they do.

Fourth, the approximate solution of the differential equation must be computed quickly and must be smooth enough (as a function of the parameters) to be used in empirical work. Finite difference approximations to the equilibrium price function do not meet these criteria. In particular, since the differential equation is defined implicitly at each point in the state space, finite difference methods require the solution of a non-linear equation at each step of the integration. I develop a projection method approximation that is many times faster.[4]This method approximates the solution of the ODE with a piecewise polynomial that solves the ODE at a set of optimally chosen points in the state space.

The result of these computations is an equilibrium price function and a set of functions describing the distributions of education, income, preference, and ability within each location. I trace several fundamental ways these equilibrium functions differ from the equilibrium solutions of the linear utility model. In general, they demonstrate that the model developed in section 5 can generate much more varied patterns of sorting than can be generated by the linear utility model. These results cannot be obtained without developing the computational tools discussed above.

Equipped with these computational tools, I analyze the empirical equations resulting from the exponential utility model in section 5.4. The set of empirical equations is very similar to the set of equations resulting from the linear utility model. Both the hedonic location choice equation and the education production function contain the same elements as in section 4.3. Now, however, the logarithm of the marginal price and the logarithm of the quality variable (the two endogenous variables that enter the location choice equation) are not collinear. I show that this non-linearity is sufficient to identify the parameters of the location choice equation up to scale. Moreover, I show that this is also sufficient to identify the parameters of the education production function.

In section 5.5, I examine data generated from the equilibrium model and find that maximum likelihood estimation recovers estimates of the structural parameters with a high degree of precision. Very few of the estimates reject the hypothesis that the true parameters are the values used to generate the synthetic data. I also present preliminary empirical results analyzing the NELS dataset. These estimates demonstrate the efficacy of the methods developed in this paper. Future work will analyze alternative functional forms and more flexible functional forms that can more accurately approximate the

relationships observed in the data.

3 Model

There is a distribution of consumers in the economy each of whom chooses a residential location based on the qualities of schools that are available in different locations. Each consumer is characterized by a vector of traits that affect the utility they obtain from schools of various qualities. I represent these traits by the vector $x = (\ln s_0, \ln w_0, \ln a, \ln \beta, x_5)'$ where $s_0 = e^{x_1}$ is parental education or schooling attainment, $w_0 = e^{x_2}$ is parental income or wealth, $a = e^{x_3}$ is the ability of the consumer's child in school, $\beta = e^{x_4}$ is a preference parameter that measures how much the consumer cares about their child's schooling relative to their own consumption, and x_5 is a shock to the child's educational process that is realized after the parent makes their choice of location. Throughout the paper, I will use x_i to indicate the i'th component of x and will use $X = R_5$ to indicate the set of all consumers.

I assume that the population distribution of the consumer characteristics is log-normal, i.e. $x \sim N(\mu, \Sigma)$. The distribution of consumer types is one of the crucial determinants of the shape of the equilibrium because it determines

both the relative demand for different locations in the economy as well as the relative supply of different types of consumers.

The consumers choose their residential neighborhoods from a continuum of locations indexed by $z \in R_+$. Each location contains an inelastic supply density of indivisible residential houses $h(z)$ owned by competitive landlords who rent to the consumers at the price $q(z)$ per unit.[5] In addition, each location is characterized by the set of residents who live there. This is determined by the measurable equilibrium assignment function $F(x) : X \to R_+$ that assigns each person x to a location z. The set of all residents in each location is $\{x : x \in F^{-1}(z)\}$ where $F^{-1}(z)$ is the preimage of the set $\{z\}$. These residents determine the quality of the schools in location z. In particular, I assume that the quality of the schools in neighborhood z is determined by the average schooling of the people living in z. Let $S(z)$ represent the average schooling in location z and define $S(z) = E[e^{x_1} | x \in F^{-1}(z)]$. Note that the population living in any set of locations $Z \in \mathcal{B}$, where $\mathcal{B}$ is the Borel

[5] I assume that $h(z)$ is a positive continuous density function and that $\int h(z)\, dz \geq 1$. This ensures that the total supply of houses is sufficient to house the population of consumers and that there are no neighborhoods that have a positive point mass of housing supply.

σ-algebra on R_+, is then given by

$$P(Z) = \int_{F^{-1}(Z)} \phi_5(x, \mu, \Sigma)\, dx$$

where $\phi_5(x, \mu, \Sigma)$ is the five-dimensional normal probability density function with mean μ and variance Σ . Further note that in equilibrium, this population measure can have no mass points at any location since $h(z)$ has no mass points. Therefore, the equilibrium population measure can be represented as $P(Z) = \int_Z dP(z)\, dz$ where $dP(z)$ is the Radon-Nikodym derivative of the measure P with respect to the Borel measure on R_+.

School quality is important to consumers because it interacts with family background characteristics to produce childrens' educational outcomes. I assume that a child's educational outcome, which I denote by s_1 for schooling attainment, is a function of four inputs. It is a function of the average parental schooling attainment in the neighborhood $S(z)$, parental schooling attainment $s_0 = e^{x_1}$, the child's ability $a = e^{x_3}$, and the shock x_5.

$$s_1 = S(z)^{\eta_1} e^{\eta_2 x_1 + x_3 + x_5} \tag{3.1}$$

I assume that when a parent chooses their residential location, they know

the quality of the neighborhood S, their own education e^{x_1}, and their child's ability e^{x_3}. These variables affect their residential decision. They do not know the value of the variable x_5. They only know the distribution from which it is drawn. Income, $w_0 = e^{x_2}$ and the preference parameter $\beta = e^{x_4}$ do not directly affect the educational production function but may indirectly affect educational outcomes through their affect on location choice.

Given the above information, I can state the consumer's utility maximization problem. They treat the price function $q(z)$ and the distribution of types across neighborhoods $F(x)$ as given and solve:

$$\max_z \left\{ \frac{1 - e^{-\gamma(e^{x_2} - q(z))}}{\gamma} + e^{x_4} E(s_1|z) \right\} \tag{3.2}$$

where $E(s_1|z) = A_0 S(z)^{\eta_1} e^{\eta_2 x_1 + x_3}$, $S(z) = E[e^{x_1}|x \in F^{-1}(z)]$,and $A_0 = E(e^{x_5})$.

Utility is a function of consumption, $e^{x_2} - q(z)$, and the expected value of the child's schooling attainment, $E(s_1|z)$. The preference parameter e^{x_4} measures the weight a parent puts on their child's expected schooling.[6]

The maximization problem in (3.2) describes the utility function for one

[6] e^{x_4} has two interpretations. It reflects parental altruism and parental perceptions about their child's expected return to education measured in utility units.

consumer with a fixed set of characteristics, $x = (\ln s_0, \ln w_0, \ln a, \ln \beta, x_5)'$. Its solution determines the demand correspondence for that individual consumer, $Z = d(x, F, q)$. Z is the set of locations that maximizes the utility of the consumer with characteristics x when the distribution of consumers across locations is described by the function F and the price schedule is q. Using these objects, I can now define a locational equilibrium.

Definition 3.1 *Let $d(x, F, q)$ be the solution of (3.2) for consumer x. A locational equilibrium of this economy is a pair of measurable functions (F, q) such that $F(x) : X \longrightarrow R_+$, $q(z) : R_+ \longrightarrow R_+$, and*

1. *if $x \in F^{-1}(z)$, then $z \in d(x, F, q)$ for all $x \in X$*
2. $\int_Z (h(z) - dP(z))\, dz \geq 0$ *for all $Z \in \mathcal{B}$*
3. $\int_Z q(z)(h(z) - dP(z))\, dz = 0$ *for all $Z \in \mathcal{B}$*
4. $q(z) \geq 0$ *for all $z \in R_+$*

The first condition requires that an equilibrium assignment of consumers to neighborhoods assigns each consumer to one of his optimal locations. The second and third conditions require that housing demand is never larger than housing supply and equals housing supply in every neighborhood that has a

positive price. The fourth condition requires that landlords profits are non-negative. The equilibrium is a variation of a Nash equilibrium in that each agent assumes that every other agent's equilibrium choice is fixed when he chooses his own optimal action. The Nash equilibrium consistency condition that these assumptions are correct is then imposed by condition 1) in the above definition.

The immediate question arises, under what conditions does an equilibrium exist. A simple answer, is that a trivial equilibrium always exists. This answer is given in the following theorem.

Theorem 3.1 *Let* $\overline{S} = E\left[e^{x_1}\right]$ *be the average education in the economy. A locational equilibrium* (F, q) *exists that satisfies:*

$$E\left[e^{x_1} \left| x \in F^{-1}(z)\right.\right] = \overline{S} \quad \forall z \in R_+$$

$$q(z) = 0 \qquad \forall z \in R_+$$

Proof. See appendix A. ■

In this equilibrium every location has the same quality and the same price. This trivial equilibrium is not very interesting. More interesting are separating equilibria in which different locations have different qualities and prices. However, when these equilibria exist, they are not unique. Many

different equilibrium assignments of consumers to locations are possible.

Lemma 3.2 *Let* (F, q) *be a locational equilibrium. Let* $\left(\widehat{F}, \widehat{q}\right)$ *be a pair that satisfies conditions 2-4 in definition* (3.1). $\left(\widehat{F}, \widehat{q}\right)$ *is a locational equilibrium if it satisfies:*

1. *For all* $x' \in X$, *if* $x' \in F^{-1}(z_1)$, $S_1 = E\left(e^{x_1} | x \in F^{-1}(z_1)\right)$, *and* $x' \in \widehat{F}^{-1}(z_2)$, *then* $S_1 = E\left(e^{x_1} | x \in \widehat{F}^{-1}(z_2)\right)$

2. *2) For all* $z_1, z_2 \in R_+$, *if*

$$S_1 = E\left(e^{x_1} | x \in F^{-1}(z_1)\right) and S_1 = E\left(e^{x_1} | x \in \widehat{F}^{-1}(z_2)\right)$$

then $q(z_1) = \widehat{q}(z_2)$.

Proof. See appendix A. ■

If (F, q) is a locational equilibrium, any pair $\left(\widehat{F}, \widehat{q}\right)$that changes consumers' locations in z-space without changing their location in $(S, p(S))$space is also a locational equilibrium if it maintains equilibrium in the housing supply market in every location. This type of reassignment can change the price and quality at a particular location z but does not change the welfare of any consumers. It amounts to a redistribution of wealth among landlords in different locations.

All equilibria satisfying lemma (3.2) share two traits. They maintain the same quality price schedule and they assign the same sets of consumers to each neighborhood quality level . In this sense, they belong to a unique class of equilibria described by a common price schedule that assigns a unique price to each neighborhood quality and by a common rule assigning consumers to different quality neighborhoods. I call such a class of equilibria a "quality sorting equilibrium class" or simply a "quality sorting equilibrium."

Definition 3.2 *A quality sorting equilibrium is a pair of measurable functions* (G, p) *such that* $G : X \to R_+$, $p : G(X) \to R_+$, *and*

1. $E\left[e^{x_1} \left| x \in G^{-1}(S)\right.\right] = S$, *for all* $S \in G(X)$

2. *if* $G(x) = S$, *then* $S \in \underset{S' \in G(X)}{\arg\max} U\left(w_0 - p\left(S'\right), S', x\right)$ *for all* $x \in X$

In the rest of this dissertation, I do not distinguish between equilibria within a quality sorting equilibrium class since all equilibria in this class have identical implications for consumers who are the primary subjects of this study. Instead, I focus the analysis on quality sorting equilibria. This focus is valid since, except in pathological examples, I can always find a location sorting equilibrium consistent with a given quality sorting equilibrium.

Theorem 3.3 *Let $\mathcal{B}$ be the Borel σ-algebra on R_+. Fix a quality sorting equilibrium (G, p). Using G define the measure μ on $\mathcal{B}$ via*

$$\mu(\mathcal{S}) = \int_{G^{-1}(\mathcal{S})} \phi_5(x)\, dx$$

for all $\mathcal{S} \in \mathcal{B}$. If $\mu = \mu_a + \mu_d$, where μ_a is an absolutely continuous measure and μ_d is a discrete measure, then there is a locational sorting equilibrium (F, q) that implements (G, p) such that

1. *$G(x) = S$ if and only if $S(F(x)) = S$ for all x*

2. *if $S = E[e^{x_1} | x \in F^{-1}(z)]$, then $q(z) = p(S)$ for all z*

Proof. See appendix A. ■

The questions of existence and uniqueness of a separating quality sorting equilibrium are more complicated. They depend more closely on the particular parameters of the model studied. In sections 4 and 5, I study examples in which a separating equilibrium exists. In both these examples, utility is an exponential function of consumption. In section 4, I examine the simpler limiting case in which γ, the coefficient of absolute risk aversion in the utility function, is zero. In this limiting case, utility is linear in consumption and there is a separating equilibrium if and only if the correlation between

parental education and the parental willingness to pay for education is positive. In this case, people with more education are willing to pay more on average for high quality locations. This differential willingness to pay is sufficient to sustain a separating equilibrium with higher quality neighborhoods having higher average education.

In section 5, I examine the more complicated case in which $\gamma > 0$ and the exponential utility model does not reduce to a linear model. In this case, no closed form analysis of the equilibrium is possible and exact conditions necessary for the existence of a separating equilibrium are not available. Instead, I analyze cases where conditions sufficient for the existence of a separating equilibrium are satisfied.

4 Linear utility model

Here I develop a specific example of the general model described in section 3 and analyze its equilibrium. The equilibrium is a quality sorting equilibrium as defined in definition 3.2. In the example, I impose that γ, the coefficient of absolute risk aversion, is zero. This allows me to find a unique closed form solution for the quality price function and for the conditional distributions of

consumer types. The development of this result introduces many of the ideas that are used in the analysis of the more complicated model in section 5. I derive the price function and related results in sections 4.1 and 4.2. Then after analyzing the results, I examine their empirical implications in section 4.3.

4.1 Equilibrium

Finding an equilibrium requires solving the following problem.

Problem 4.1 *Let $\mathcal{G}$ be the set of measurable functions on X and let $\mathcal{P} = C^2(R_{++}) \bigcap C^0(R_+)$ [7] be the set of twice continuously differentiable functions. The problem is to find the pair $(G, p) \in \mathcal{G} \times \mathcal{P}$ that satisfy:*

1. $E[e^{x_1} | x \in G^{-1}(S)] = S$ for all $S \in G(X)$

2. For all $x \in X$, if $S = G(x)$, then

$$S \in \underset{S' \in G(X)}{\arg\max} \left\{ e^{x_2} - p(S') + A_0 S'^{\eta_1} e^{\eta_2 x_1 + x_3 + x_4} \right\} \tag{4.1}$$

[7] $C^2(R_{++})$ *is the space of functions that are twice continuously differentiable everywhere on the domain* R_{++}. $C^0(R_+)$ *is the space of functions that are continuous on the domain* R_+.

I solve the consumer's maximization problem and then impose conditions 1 and 2 to find the equilibrium. The consumer treats (G, p) as exogenous and solves the maximization problem (4.1). For a particular consumer described by a fixed vector of characteristics x, the solution to (4.1) describes his optimal neighborhood choice, S, through the first-order condition:

$$p_S = \eta_1 A_0 S^{\eta_1 - 1} e^{\eta_2 x_1 + x_3 + x_4} \tag{4.2}$$

where p_S denotes $\frac{dp}{dS}$.[8] More importantly, for each fixed level of S, this same first-order condition describes the set of people who choose to live in each neighborhood. This set is simply the set of all people whose vector of characteristics satisfy (4.2). Taking the logarithm of (4.2), and defining $f(S) = \ln p_S - \ln(A_0 \eta_1) + (1 - \eta_1) \ln S$ this set is

$$g(S, p_S) = \{x \,|\, f(S) = B'x\} \tag{4.3}$$

where $B = \begin{bmatrix} \eta_2 & 0 & 1 & 1 & 0 \end{bmatrix}'$.

[8]For the moment, I assume that every consumer's second-order condition is globally satisfied. After solving the equilibrium pricing equation, I will check that this assumption is true.

This expression illustrates how the sorting economy matches a one dimensional index of location quality, $f(S)$, to a one dimensional index of consumer willingness to pay (WTP), $B'x$. $f(S)$ is an index that summarizes the effect of locational quality and price on marginal utility. The consumer WTP index $B'x$ is a random variable that measures each consumer's marginal valuation of locational quality. The equilibrium partitions consumers into sets indexed by $f(S)$ by matching higher values of $f(S)$ to higher values of WTP.

In equilibrium, assuming a unique maximum for each consumer, condition 2 in problem 4.1 implies $G^{-1}(S) = g(S, p_S)$. Combining this with condition 1, the requirement that $S = E[e^{x_1} | x \in G^{-1}(S)]$ for all S, results in a differential equation describing the equilibrium price:

$$S = E[e^{x_1} | x \in g(S, p_S)] \tag{4.4}$$

An equilibrium price must satisfy (4.4) for all S.

I proceed to derive an explicit representation for the right side of equation (4.4). The definition of $g(S, p_S)$ implies that a consumer chooses location S if $f(S) = B'x$. Recalling that $x \sim N(\mu, \Sigma)$, the distribution of education

among the people who choose S is

$$(x_1 \,|f\,(S) = B'x\,) \sim N\left(\theta_s, \Psi_s^2\right) \tag{4.5}$$

where $\theta_s = \mu_1 + (B'\Sigma e_1)\,(B'\Sigma B)^{-1}\,(f\,(S) - B'\mu)$ and $\Psi_s^2 = \sigma_{11} - \frac{(B'\Sigma e_1)^2}{(B'\Sigma B)}$. μ_1 is the first component of μ and σ_{11} is the variance of x_1. In words, the distribution of parental schooling in the population within each neighborhood is log-normal with mean θ_s and variance Ψ_s^2. As a result the average education of those who live in neighborhood S is given by:

$$E\left[e^{x_1} \,|x \in g\,(S, p_S)\right] = e^{\theta_s + 0.5\Psi_s^2} \tag{4.6}$$

This average is a function of μ, the mean characteristics in the population, Σ, the covariance matrix of the population characteristics, η_1 and η_2, the parameters of the education production function, and finally S and p_S since θ_S depends on S and p_S through $f\,(S)$. Moreover, substituting this formula into (4.4) implies that an equilibrium price function must satisfy

$$S = L_0 p_S^{L_1} S^{L_1(1-\eta_1)} \tag{4.7}$$

where $L_0 = e^{\mu_1 - L_1(\ln(\eta_1 A_0) + B\mu) + 0.5\Psi_s^2}$ and $L_1 = (B\Sigma e_1')(B\Sigma B')^{-1}$.

L_0 and L_1 are constants that depend on the distribution of population characteristics. L_0 determines the level of the price premium consumers pay to live in higher quality neighborhoods. L_1 is the regression coefficient from the regression of log-education on the WTP for neighborhood quality. It determines the curvature of the price function.

Despite the fact that the price function must satisfy each person's first-order condition (4.2), equilibrium consistency and the population distribution of types of people impose the restriction that the price function must identically satisfy the ordinary differential equation (4.7). This equation defines a unique family of price functions that are consistent with equilibrium. Adding the initial condition provided by the consideration that zero quality neighborhoods must have a price of zero, pins down the solution.[9]

$$p(S) = \frac{S^{\eta_1 + \frac{1}{L_1}}}{(L_0)^{\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1}\right)} \tag{4.8}$$

[9]The price function is convex if $\eta_1 + \frac{1}{L_1} \geq 1$. The consumer second-order condition for maximization is globally satisfied for every consumer as long as $L_1 > 0$ and $\eta_1 > 0$. These conditions are satisfied if people with more education pay more for school quality on average and if the marginal product of S is positive.

The constants in (4.8) are functions of the parameters of the population distribution of characteristics and of the parameters of the education production function. η_1 and L_1 play primary roles. They are the sole determinants of the elasticity of the price premium with respect to neighborhood quality. η_1 measures the elasticity of children's schooling attainment with respect to school quality. L_1 measures the correlation between willingness to pay for school quality and parents' own schooling attainment. Thus, a sorting equilibrium maps the importance of school quality in the production of children's education and the degree of correlation between parental education and willingness to pay for neighborhood quality directly into percentage price differences across neighborhoods. When η_1 is large, school quality is highly important for educational outcomes. Large percentage price differences across neighborhoods are required to segregate people into their preferred locations. Similarly, when the degree of correlation between parental education and willingness to pay for neighborhood quality is high (so that L_1 is large), small price differentials are required to maintain the equilibrium segregation of people because people with similar willingness to pay are relatively homogenous in education.

4.2 Conditional distributions of consumer types

The model also predicts the equilibrium within-neighborhood distribution of population characteristics. Since the predictions are analagous for all characteristics, I only discuss log-schooling and log-ability in detail. Consider the conditional distribution of log-schooling, $x_1 = \ln s_0$. After substituting the equilibrium price function into (4.5) , one can see that

$$(x_1 \,|S\,) \sim N\left(\theta_s, \Psi_s^2\right)$$

where $\theta_s = \ln S - 0.5\Psi_s^2$ and $\Psi_s^2 = \sigma_{11}\left(1-\rho_1^2\right)$. σ_{11} is the population variance of log-education. ρ_1 is the correlation between log-education and WTP for neighborhood quality.[10]

Ψ_s^2 , the conditional variance of log-education is constant across neighborhoods and is smaller than the population variance. Since people sort based on common willingness to pay and since that willingness to pay is correlated with parental education, individual neighborhoods are more homogenous in terms of education than the population at large. How much more homogenous depends on ρ_1, the correlation between log-education and

[10]This formula for Ψ_s^2 is equivalent to that in (4.5).

WTP. The larger is ρ_1, the smaller is the within-neighborhood variance of log-education.[11]

For instance, if Σ is a diagonal matrix and σ_{ii} is the (i,i) component of Σ, then

$$\rho_1^2 = \frac{\eta_2^2 \sigma_{11}}{\eta_2^2 \sigma_{11} + \sigma_{33} + \sigma_{44}}$$

Thus, ρ_1^2 is large when the product of η_2 and the population variance of log-schooling (σ_{11}) is large relative to the population variance of log-ability (σ_{33}) and log-preference (σ_{44}) so that the variance of log-schooling, is the predominant component of the variance of WTP in the population. In the limit, as η_2 or σ_{11} approach infinity, all of the variance of WTP is due to the variance of log-education, $\rho_1^2 \to 1$, and the within-neighborhood variance of log-schooling approaches zero as sorting based on willingness to pay is equivalent to sorting based on parental education.

Similar results apply to the conditional distributions of log-ability, log-preference, and log-income. Since all of these characteristics are components of x, their distributions conditional on $A(S) + B'x = 0$ are also normal.

[11] Using the National Educational Longitudinal Survey of 1988 (NELS), I find that the population variance of log-education is 0.035 while its within-school variance averages 0.023. This implies a value for ρ_1 of 0.59.

Recalling that $x_3 = \ln a$, the conditional distribution of log-ability is:

$$(x_3 \,|S\,) \sim N\left(\theta_a, \Psi_a^2\right)$$

where $\theta_a = \mu_3 + \frac{\sqrt{\sigma_{33}}\rho_3}{\sqrt{\sigma_{11}}\rho_1}\left(\ln S - \mu_1 - 0.5\Psi_s^2\right)$ and $\Psi_a^2 = \sigma_{33}\left(1-\rho_3^2\right)$. σ_{33} is the variance of log-ability and ρ_3 is the correlation between log-ability and WTP for neighborhood quality. Notice that while one might expect ρ_3 to be positive it could be positive or negative.

As with log-schooling, the correlation between log-ability and WTP determines the ratio of the within-neighborhood variance of log-ability to the population variance. The larger ρ_3^2, the smaller the ratio. Also, holding other things constant, increases in the variance of log-ability increase the correlation of log-ability and WTP. This follows because these increases make log-ability a larger and larger component of the variance of WTP. In the limit, when log-ability and WTP are highly correlated, differences in WTP which separate people in the housing market in essence separate them in terms of ability.

These formulas also indicate how mean log-ability varies with neighborhood quality. It varies linearly and is strictly increasing with neighborhood

quality if $\rho_3 > 0$ and $\rho_1 > 0$.[12] The more important ability is as a component of the variance of, the more rapidly mean log-ability increases with neighborhood quality. In the diagonal covariance matrix case

$$\frac{\sqrt{\sigma_{33}}\rho_3}{\sqrt{\sigma_{11}}\rho_1} = \frac{\sigma_{33}}{\eta_2\sigma_{11}}$$

Thus, when σ_{33} is large or when $\eta_2\sigma_{11}$ is small the conditional mean of log-ability has a large slope. Intuitively, these facts illustrate the point that when variation in WTP is more closely related to variation in log-ability than to variation in log-education, there will be larger differences in mean abilities across neighborhoods than when the reverse is true.

The conditional distributions of log-preference and log-income are analogous. For each trait, the key parameter governing the elasticity of the conditional mean with respect to locational quality is the correlation of that trait with WTP.[13]

[12] While one might expect $\rho_3 > 0$, it could be negative. For instance, if parents of low ability children value children's educational outcomes very highly then ρ_3 could be negative.

[13] Though income plays no direct role in sorting, to the extent that it is correlated with the other variables, its conditional distribution is a non-trivial function of neighborhood quality.

4.3 Empirical implications

The theoretical model from section 4 yields three empirical equations: the education production function, (3.1) , the consumer first-order condition from the neighborhood choice problem, (4.2) , and the equilibrium neighborhood price equation, (4.8). Since the error terms in these equations (log-ability and log-preference) can be freely correlated with log-education and log-income, before rewriting the system, I decompose the unobservables into components that are correlated with log-income and log-education and components that are uncorrelated with these observable variables. I reparameterize the unobservables as follows:

$$\begin{aligned} x_3 &= \alpha_1 + \alpha_2 x_1 + \alpha_3 x_2 + \varepsilon_3 \\ x_4 &= \beta_1 + \beta_2 x_1 + \beta_3 x_2 + \varepsilon_4 \end{aligned} \tag{4.9}$$

Hence, ε_3 and ε_4 are the components of log-ability and log-preference that are uncorrelated with log-education and log-income. Using these representations

of the unobervables, the system of econometric equations is

$$\begin{aligned} \ln s_1 &= \alpha_1 + \eta_1 \ln S + \widehat{\alpha}_2 x_1 + \alpha_3 x_2 + z_1 \qquad (4.10) \\ \ln p_S + (1-\eta_1)\ln S &= \alpha_1 + \beta_1 + \ln(\eta_1 A_0) + \widehat{\eta}_2 x_1 + \widehat{\eta}_3 x_2 + z_2 \\ p(S) &= \frac{S^{\eta_1 + \frac{1}{L_1}}}{(L_0)^{\frac{1}{L_1}}\left(\eta_1 + \frac{1}{L_1}\right)} \end{aligned}$$

where $\widehat{\alpha}_2 = \eta_2 + \alpha_2$, $\widehat{\eta}_2 = \eta_2 + \alpha_2 + \beta_2$, $\widehat{\eta}_3 = \alpha_3 + \beta_3$, $z_1 = \varepsilon_3 + x_5$, and $z_2 = \varepsilon_3 + \varepsilon_4$.

x_1 and x_2 are the observable levels of log-education and log-income and z_1 and z_2 are the components of the error terms that are orthogonal to the observables. By assumption z_1 and z_2 are jointly normally distributed with covariance matrix

$$\Gamma = \begin{bmatrix} \sigma_{33} + \sigma_{55} & \sigma_{33} + \sigma_{34} \\ \sigma_{33} + \sigma_{34} & \sigma_{33} + \sigma_{44} + 2\sigma_{34} \end{bmatrix}.$$

The parameters I would like to estimate are the education production function parameters (η_1 and η_2) and the parameters describing the distribution of consumer characteristics $(\alpha_i, \beta_i, \Gamma)$ for $i = 1, 2, 3$. Notice immediately that while it is possible that $\widehat{\alpha}_2 = \eta_2 + \alpha_2$ can be identified, η_2 and α_2 cannot be

independently estimated since they do not have linearly independent effects on the outcome. A further instrument is needed to disentangle these two effects. This however, is not the primary issue raised by the location choice demand system. I will focus the remainder of the discussion on identification of the remaining structural parameters.

The first equation in the system (4.10) is the typical equation estimated in the neighborhood effects literature. It relates the child's educational outcome to school quality, observable parental data, and unobserved child and family traits. Since $\ln S$ is correlated with z_2, it is immediate that estimation of this equation alone cannot identify the structural parameters unless z_1 and z_2 are uncorrelated; that is unless $Var(\varepsilon_3) = -cov(\varepsilon_3, \varepsilon_4)$. This condition can only be met if changes in the conditional mean log-preference across neighborhoods exactly offset changes in the conditional mean log-ability. If this condition is not met,then estimation of the education production function will produce biased parameter estimates. Moreover, the direction of the bias is unknown since the bias depends on the covariance between $\ln S$ and log-ability. As noted in the previous section, this covariance can be either positive or negative.

The second and third equations in (4.10) are the hedonic demand equa-

tion and the equilibrium price function. The classical approach to hedonic estimation attempts to estimate these two equations by first fitting an arbitrary functional form to data on prices and characteristics of locations, and then by using this fitted marginal price function in a second stage estimation of the hedonic demand equation in (4.10). It is immediate that this approach cannot succeed in this economy because equilibrium requires that $\ln p_S$ is collinear with $\ln S$.

Next consider estimation of the system. After substituting the price function into the hedonic demand equation, the system (4.10) reduces to a set of linear simultaneous equations

$$\begin{aligned} \ln s_1 &= \alpha_1 + \eta_1 \ln S + \widehat{\alpha}_2 x_1 + \alpha_3 x_2 + z_1 \\ \ln S &= L_1\widehat{\eta}_0 + L_1\widehat{\eta}_2 x_1 + L_1\widehat{\eta}_3 x_2 + L_1 z_2 \end{aligned} \tag{4.11}$$

where $\widehat{\eta}_0 = -B\mu + \mu_s + 0.5\Psi_s^2 + (\alpha_1 + \beta_1)$. The first equation of the system is not identified unless z_1 and z_2 are uncorrelated or $\alpha_3 = 0$ and $\beta_3 \neq 0$.[14] Even then, the structural parameters in the second equation are not identified. In particular, L_1, which measures the population correlation between parental

[14] Alternatively, it can be identified in a classical manner if another instrument that determines location choice without affecting the production of education is available.

education and WTP for school quality, cannot be separated from the $\widehat{\eta}_i$ which measure the direct effects of parental characteristics on location choice. This failure of identification is a direct result of the equilibrium collinearity between $\ln p_S$ and $\ln S$.

Nevertheless, as is clear from (4.11) , certain combinations of the parameters can be estimated with data on education, income, school quality, and educational outcomes. Tables 1-5 display estimates of these combinations of parameters obtained from an analysis of the National Educational Longitudinal Survey (1988) assuming the linear utility model is the true model of locational choice and educational production. Table 1 displays estimates of the first equation in (4.11) alone. If z_1 and z_2 are uncorrelated, then these estimates imply that the elasticity of children's education with respect to school quality is 0.328 (0.0302) and the elasticity with respect to parental education is 0.376 (0.0189). Table 2 displays estimates of the second equation in (4.11) alone. Education and income explain about 20% of the variation in location choice.

Estimates of the reduced form of the system (4.11) are displayed in Table 3. Since these parameter estimates combine the effects of several of the structural parameters, the values estimated are hard to interpret. However,

Tables 4 and 5, use these estimates to explore implications these estimates have for values of the structural parameters. Table 4 shows the parameter estimates that are obtained if I assume that the system is identified, i.e. if I assume that $\alpha_3 = 0$. This assumption states that the partial correlation coefficient between income and ability is zero. If this is true, then Table 4 shows that η_1, the elasticity of children's education with respect to school quality, is 0.540 (0.0936) and $\widehat{\alpha}_2$, the elasticity of children's education with respect to their own parent's education (controlling for the correlation between parental education and child's ability), is -0.0222 (0.0172).

These results depend on the assumption that $\alpha_3 = 0$. Table 5 investigates how these results depend on the value of α_3. If α_3 is greater than about 0.007, then the estimate of η_1 is negative. Therefore, I restrict the table to values of $\alpha_3 < 0.007$. The smaller is α_3, the larger is η_1 and the smaller is $\widehat{\alpha}_2$. Nearly any values of the parameters are possible. Given this model, it is impossible to infer whether parental background characteristics or school quality are important determinants of children's educational outcomes. This is empirical confirmation that the equlibrium collinearity between $\ln p_S$ and $\ln S$ prevents identification of the model.

5 A sorting economy with exponential utility

In deriving the equilibrium conditions in section 4, I showed that the two determinants of the shape of the equilibrium pricing function are the individual consumer demand functions (as determined by the utility function and the educational production function) and the distribution of consumer characteristics in the population. In the example solved, I made assumptions about these two factors that lead to a closed form solution. However, small changes in the assumptions made about the utility function or the distribution of consumer types lead to an equilibrium that does not have a closed form solution. They also lead to an equilibrium that has more varied patterns of sorting and pricing. These characteristics of an equilibrium of models perturbed away from the un-identified state, lead to equilibria whose empirical systems of equations are identified.

The simplest extension of the linear utility model that yields these results relaxes the assumption that utility is linear in consumption. This generalization has three important benefits. Most simply, it allows income to play a non-trivial role in sorting. Secondly, it results in an equilibrium price function that is not a constant elasticity function of neighborhood quality. This increases the empirical power of the model because it destroys the linearity

that prevented the model in section 4 from identifying the parameters. Finally, this extension results in much richer patterns of sorting in the economy.

5.1 Equilibrium with exponential utility

The conceptual approach is analogous to that in section 4. As before there is a continuum of neighborhoods indexed by quality and a continuum of consumers indexed by x. Consumers treat prices and location qualities as fixed, and choose their optimal locations. I characterize the set of consumers who choose to live in each location, impose a Nash equilibrium consistency condition, and solve the ordinary differential equation that matches each consumer to a set of consumers with the same willingness to pay for school quality.

The problem is slightly more complex than the linear utility model for two main reasons. First, numerical methods must be used to calculate the average education in each location and to solve the equilibrium pricing equation. Second, I must allow for price functions that have a kink. For some values of the parameters, a kinked price function is required to ensure that all consumers second-order conditions are satisfied. Despite these complications, I am able to derive conditions characterizing equilibrium in the model,

compute and analyze the equilibrium of the model, and use these results to show that the system of empirical equations resulting from the equilibrium can be used to estimate the structural parameters in the model.

The problem to be solved is the following.

Problem 5.1 *Let $\mathcal{G}$ be the set of measurable functions on $X = R_5$, let $\mathcal{S}$ be a finite set of points in R_{++}, and let $\mathcal{P} = C^2(R_{++} \setminus \mathcal{S}) \bigcap C^0(R_+)$ be the set of piecewise twice continuously differentiable functions on R_+. Noting that $s_0 = e^{x_1}$ and $w_0 = e^{x_2}$, the problem is to find the pair $(G, p) \in \mathcal{G} \times \mathcal{P}$ that satisfy:*

1. $E[e^{x_1} | x \in G^{-1}(S)] = S$ for all $S \in G(X)$

2. For all $x \in X$, if $x \in G^{-1}(S)$, then

$$S \in \underset{S' \in G(X)}{\arg\max} \left\{ \frac{1 - e^{-\gamma(e^{x_2} - p(S))}}{\gamma} + A_0 S'^{\eta_1} e^{\eta_2 x_1 + x_3 + x_4} \right\} \tag{5.1}$$

The equilibrium definition must allow for piecewise twice continuously differentiable equilibrium price functions because for some parameter values, no twice continuously differentiable function satisfies the equilbrium conditions. I will discuss this point more in section 5.2.

First, I characterize the set of people who choose each location. Given a pair (G, p), each consumer chooses an optimal location S. The consumer first-order condition is

$$-e^{-\gamma(e^{x_2}-p)}p_S + \eta_1 A_0 S^{\eta_1 - 1} e^{\eta_2 x_1 + x_3 + x_4} = 0 \tag{5.2}$$

Assuming that the second-order condition is globally satisfied for all consumers, the set of people satisfying (5.2) is the set who choose location S. Taking logarithms and defining

$$f(S, p, p_S) = (1 - \eta_1) \ln S + \gamma p + \ln p_S - \ln(A_0 \eta_1) \tag{5.3}$$

this set can be written

$$g(S, p, p_S) = \{x \,|\, f(S, p, p_S) = \eta_2 x_1 + \gamma e^{x_2} + x_3 + x_4\} \tag{5.4}$$

As in the previous section, this expression illustrates how the sorting economy matches a one dimensional index of location quality, $f(S, p, p_S)$, to a one dimensional index of consumer willingness to pay (WTP), $\eta_2 x_1 + \gamma e^{x_2} + x_3 + x_4$. $f(S, p, p_S)$ is an index that summarizes the effect of locational

quality and price on marginal utility. The consumer index of WTP (the right side of (5.4)) is a random variable that measures peoples' willingness to pay (WTP) for neighborhood quality. The equilibrium partitions consumers into sets indexed by f. This index plays a crucial role throughout the subsequent analysis. In the theoretical section, it enables me to prove several facts about the equilibrium price function. In the empirical section, it plays the role of the choice variable in the consumer's reduced form locational choice equation.

Given $g(S, p, p_S)$, the set of people who choose to live in each location S, condition 2 of problem 5.1 requires that $g(S, p, p_S) = G^{-1}(S)$. Combining this with condition 1 of problem 5.1 leads to the following differential equation characterizing equilibrium

$$S = E\left[e^{x_1} \,|x \in g(S, p, p_S)\right] \tag{5.5}$$

with initial condition $p(0) = 0$. This condition is analogous to the equilibrium condition in the linear model. In contrast to the linear model, however, this differential equation does not have an analytical solution. Instead, as detailed

in appendix B, it reduces to an equation of the form

$$S - F\left(f\left(S, p, p_S\right)\right) = 0 \tag{5.6}$$

where $f\left(S, p, p_S\right)$ is defined in equation (5.3),

$$F\left(f\right) = \frac{e^{\mu_1 + \frac{1}{2}\sigma_{11}} \int e^{-\frac{1}{2}\frac{\left(z_2 - \mu_2 - \sigma_{12}\right)^2}{\sigma_{22}} - \frac{1}{2}\frac{\left(q_3(f) - \tau_{13}\right)^2}{\tau_{33}}} dz_2}{\int e^{-\frac{1}{2}\frac{\left(z_2 - \mu_2\right)^2}{\sigma_{22}} - \frac{1}{2}\frac{q_3(f)^2}{\tau_{33}}} dz_2} \tag{5.7}$$

and where $q_3\left(f\right) = f - \eta_2\xi_1 - \xi_3 - \xi_4 - \gamma e^{z_2}$.[15] $F\left(f\right)$ is the average education of the people who choose a location with quality index $f\left(S, p, p_S\right)$. An equivalent way to write it is

$$F\left(f\right) = E\left(e^{x_1} \left| WTP = f\right.\right)$$

where $WTP = \eta_2 x_1 + \gamma e^{x_2} + x_3 + x_4$. I analyze this equation and the associated equilibrium further in the next section.

There are four non-standard and non-trivial difficulties involved in analyzing equations (5.6) and (5.7). First, the integrals appearing in equation (5.7) cannot be accurately computed without careful analysis. The integrands are

[15] The parameters in equation (5.7) and in the definition of $q_3\left(f\right)$ are defined in appendix B.

highly concentrated and have extremely steep peaks. I solve this problem by developing a change of variables formula that maps the integrand into a less concentrated integrand without steep peaks that can be accurately (and quickly) approximated by Gauss-Chebyshev quadrature. Details are given in appendix C.

Second, the technique described above fails when f is either very small or very large. In both cases, numerical underflow is a problem; $F(f)$ cannot be directly evaluated because neither the numerator nor the denominator can be distinquished from zero by a finite precision computer. I solve this difficulty by deriving a Laplace type approximation for each integral and then computing the ratio of these two approximations.[16] The ratio of these two approximations can be computed for all f. Details of these approximations are given in appendix D.

Third, equation (5.6) can have singularities at points where $\frac{dF(f)}{df} = 0$. Near such points standard finite difference algorithms fail to converge to the solution of the differential equation. Fourth, related to the problem just described, equation (5.6) can have multiple local solutions. These multiple solutions make calculating an approximate global solution difficult because

[16]See Judd (1998) chapter 15 for a recent discussion on computing Laplace approximations.

any successful numerical algorithm must avoid converging to a solution that is not a global solution. For some parameter values these last two problems do not arise since $\frac{dF(f)}{df} > 0$ for all f. In these cases, the equilibrium price function is twice continuously differentiable, every consumer's second-order condition is globally satisfied, and the equilibrium population density has no mass points in quality space. In these cases the equilibrium price function can be reliably approximated with a standard numerical approximation technique. For other parameter values, however, $\frac{dF(f)}{df} \leq 0$ for some values of f. In these cases, there exists no $p \in C^2(R_{++}) \bigcap C^0(R_+)$ that satisfies the differential equation for all S. I show below that in these cases, the derivative of the equilibrium price function must have at least one discontinuity point. Moreover, positive masses of people locate at the discontinuity points. Standard approximation techniques fail because they fail to recognize that the true solution has a discontinuous first derivative. I develop an algorithm that recognizes when the true solution must have discontinuous derivative, finds the points of discontinuity, and then approximates the equilibrium price function piecewise using standard numerical techniques on subdomains where the marginal price function is continuous and imposing continuity of the solution at the discontinuity points.

5.2 Further analysis of the equilibrium equation in the model with exponential utility

The equilibrium pricing equation is given in (5.6) . It is not clear that a unique solution exists. When $\frac{dF(f)}{df} > 0$ for all f, the following theorem proves that a unique separating equilibrium exists.

Theorem 5.1 *If* $\frac{dF(f)}{df} > 0$ *for all* $f \in R$, *then there is a unique non-trivial equilibrium pricing function* $p \in C^2(R_{++}) \bigcap C^0(R_+)$.

Proof. The proof demonstrates that the equilibrium equation is equivalent to an ordinary differential equation with a unique solution. Moreover, given this solution every consumer's maximization problem has a unique solution. See appendix A for details. ■

By construction $F(f)$ is the average education of the people who have willingness to pay measured by f. The condition $\frac{dF(f)}{df} > 0$ requires that groups with higher average willingness to pay have higher average education. This guarantees that an equilibrium can be sustained in which groups with higher average education are willing to pay more for high quality locations than those with lower average education. The condition is similar to the condition $L_1 > 0$ required for sorting equilibrium in the linear model. Both

conditions imply that on average people who are willing to pay more for school quality have higher average levels of education. Thus, both conditions imply that an equilibrium can be supported in which people pay more to live in locations with higher average education.

The similarity of the two conditions can also be seen by examining $\frac{dF(f)}{df} > 0$ for small f. Lemma D.1 in appendix D shows that when $f << 0$

$$F(f) \cong L_0 e^{L_1 f} \tag{5.8}$$

where L_0 and L_1 are the constants defined in equation (4.7). When $f << 0$, $\frac{dF(f)}{df} > 0$ is equivalent to $L_1 > 0$.

Thus, $L_1 > 0$ is a necessary condition for the application of theorem 5.1. However, $L_1 > 0$ does not imply that $\frac{dF(f)}{df} > 0$ for all f. In fact, there are other examples where $\frac{dF(f)}{df} < 0$ for some f. In these examples, theorem 5.1 does not apply. However, even in these cases, there is a separating equilibrium if $L_1 > 0$.

Theorem 5.2 *If $L_1 > 0$, but $\frac{dF(f)}{df} < 0$ for some f, there is a non-trivial equilibrium pricing function $p \in C^2(R_{++} \setminus \mathcal{S}) \bigcap C^0(R_+)$ where $\mathcal{S}$ is a finite non-empty subset of R_{++}.*

Proof. See appendix A. ■

When theorem 5.2 applies, the proof in appendix A demonstrates how to construct a piecewise twice continuously differentiable price function that satisfies the equilibrium conditions. The main idea is that when $\frac{dF(f)}{df} < 0$, there is no continuous function $f(S)$ that satisfies $S - F(f) = 0$ for all S and satisfies the consumer second-order conditions for all consumers. Hence, the equilibrium index $f(S)$ must be discontinuous. This implies that the slope of the price function must be discontinuous. This further implies that a mass of people will choose to locate at the point of discontinuity in equilibrium.

5.3 Computed solutions to equilibrium pricing function

I simulate a baseline model in which $\frac{dF(f)}{df} > 0$ for all f.[17] The general shapes of the price function and its slope are shown in figures 1 and 2 for various values of γ. The $\gamma = 0$ case is displayed for comparison. When $\gamma = 0$, the price function is a constant elasticity price function. This is clearly apparent in

[17] In this section, I compute approximate solutions to the equilibrium pricing equation (5.6) using finite difference approximations. Later, in the empirical section when I must compute the solution of the equilibrium pricing equation many times, I use a projection method to approximate the price function with a piecewise polynomial approximation.

the logarithmic shaped curves in figures 1 and 2. When $\gamma > 0$, a qualitative change occurs. For low levels of S, the slope of the price function is steeper than the $\gamma = 0$ case. As S increases to moderate levels, it flattens out, before rising sharply for high values of S. When $\gamma > 0$ demand for the lowest quality neighborhoods is high relative to an economy with $\gamma = 0$. People have diminishing marginal utility of consumption. Some of them substitute towards low quality locations driving up the price of those locations relative to the $\gamma = 0$ baseline price. But, since some people shift to the low quality neighborhoods, this reduces the demand for the moderate quality neighborhoods. At these moderate levels of locational quality, average income and consumption are relatively modest and population sizes are small. Moreover, the average marginal utility of consumption is nearly constant. However, as locational quality increases diminishing marginal utility of consumption again becomes an important factor for a large fraction of the population. This results in a surge of demand leading to the sharp increase in the equilibrium prices at high values of S.

Figures 3 and 4 show the conditional means of log-income and log-ability when $\gamma = 0$ and when $\gamma > 0$.[18] For very low quality locations, all conditional

[18] $z_2 = \frac{\varepsilon_3 + \varepsilon_4}{\sqrt{\omega_{22}}}$, is the sum of the components of log-ability and log-preference that are

mean functions behave very much like the $\gamma = 0$ case. For large values of S, the conditional mean of income increases much more quickly when $\gamma > 0$. When $\gamma > 0$, income is an important determinant of location choice. On the other hand, while the conditional mean of z_2 increases with S when S is small, it declines with S when S is large. Thus, the unobserved component of ability does not increase monotonically with neighborhood quality.

Figures 5 - 7 display further characteristics of equilibrium in this model. Figure 5 shows that the variance of log-schooling initially decreases to about 95% of the baseline level before surging to 99% of the baseline level. This behavior contrasts with the flat conditional variance of log-schooling in the $\gamma = 0$ economy. Similarly, figure 6 shows that the conditional variance of log-income is sharply different in the $\gamma > 0$ economy. It reaches a peak near the mode of the quality distribution. Low and high quality locations both have low variance of log-income while medium quality locations have high variances of log-income. Figure 7 shows that the variance of z_2 falls dramatically from a peak in the lowest quality neighborhoods, reaches a minimum, and then rises again in the high quality locations.

These examples illustrate some of varied ways in which the conditional

uncorrelated with education and income. See (4.9) and (5.9) .

distributions of consumer characteristics can behave in this sorting economy. The patterns are clearly much richer than those found in equilibrium in the linear utility model. In the next section, I show that this flexibility in the patterns of sorting and in particular in the shape of the price function is sufficient for identification.

5.4 Empirical analysis of the model with exponential utility

As in section 4.3, the equilibrium yields three empirical equations: the education production function, (3.1) , the consumer's first-order condition (5.2), and the equilibrium pricing equation (5.6) . Reparameterizing the model as in equations (4.9) and substituting $f(S)$ into the consumer first-order condition, these equations are

$$\begin{aligned} \ln s_1 &= \alpha_1 + \widehat{\alpha}_2 x_1 + \alpha_3 x_2 + \eta_1 \ln S + z_1 \\ \frac{f(S)}{\sqrt{\omega_{22}}} &= \frac{\widehat{\eta}_2}{\sqrt{\omega_{22}}} x_1 + \frac{\widehat{\eta}_3}{\sqrt{\omega_{22}}} x_2 + \frac{\gamma}{\sqrt{\omega_{22}}} e^{x_2} + z_2 \\ S &= F\left(f(S), \widehat{\eta}_2, \widehat{\eta}_3, \gamma, \sqrt{\omega_{22}}\right) \end{aligned} \tag{5.9}$$

where $f(S)$ is defined in (5.3), $F(f)$ is defined in (5.7), ω_{22} is the variance of $\varepsilon_3 + \varepsilon_4$, $\widehat{\eta}_2 = \eta_2 + \alpha_2 + \beta_2$, $\widehat{\eta}_3 = \alpha_3 + \beta_3$, $\widehat{\alpha}_2 = \eta_2 + \alpha_2$, $z_1 = \varepsilon_3 + \varepsilon_5$, and $z_2 = \frac{\varepsilon_3 + \varepsilon_4}{\sqrt{\omega_{22}}}$.

x_1 and x_2 are log-schooling and log-income. They are observable. z_1 and z_2 are the unobservable error terms. By assumption z_1 and z_2 are joint normally distributed with mean zero and covariance matrix

$$\Gamma = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & 1 \end{bmatrix}$$

As in section 4.3, α_2 and η_2 cannot be separately identified; only the linear combination $\widehat{\alpha}_2 = \alpha_2 + \eta_2$ can be identified. In this section, the pricing equation implicitly defines the function $f(S)$. Importantly, this index only depends on the four parameters: $\widehat{\eta}_2$, $\widehat{\eta}_3$, γ, and ω_{22}. I want to identify these four parameters in the hedonic location choice equation as well as the parameters that enter the education production function.

First, consider the location choice equation from system (5.9). In this equation, $f\left(S, \widehat{\eta}_2, \widehat{\eta}_3, \gamma, \sqrt{\omega_{22}}\right)$ is the piecewise continuous solution to the pricing equation given by the third equation in the system. An immediate result is that this function is homogenous of degree one in the parameters.

Lemma 5.3 $f\left(S, \widehat{\eta}_2, \widehat{\eta}_3, \gamma, \sqrt{\omega_{22}}\right) = \sqrt{\omega_{22}} f\left(S, \frac{\widehat{\eta}_2}{\sqrt{\omega_{22}}}, \frac{\widehat{\eta}_3}{\sqrt{\omega_{22}}}, \frac{\gamma}{\sqrt{\omega_{22}}}, 1\right)$ *for* $\omega_{22} > 0$.

Proof. This can be checked using equations $(B.11)$ and $(B.10)$ in appendix B. ■

Thus, if the parameters in the location choice equation are identified, they are only identified up to scale and the variance of the error term in the location choice equation cannot be identified.

Consider further the remaining three parameters in the location choice equation. Abusing notation let $\widehat{\eta}_2$, $\widehat{\eta}_3$, and γ represent the values of these parameters scaled by $\sqrt{\omega_{22}}$. In terms of these scaled parameters, the likelihood for this equation is

$$\begin{aligned} \ln L\left(\widehat{\eta}_2, \widehat{\eta}_3, \gamma\right) &= -\frac{1}{2}\sum_i \left(f\left(\widehat{\eta}_2, \widehat{\eta}_3, \gamma\right) - \widehat{\eta}_2 x_1 - \widehat{\eta}_3 x_2 - \gamma e^{x_2}\right)^2 + \\ &\quad \sum_i \ln\left|\frac{df\left(S_i, \widehat{\eta}_2, \widehat{\eta}_3, \gamma\right)}{dS}\right| \end{aligned} \tag{5.10}$$

If $\gamma = 0$, the model reduces to the linear utility model in which $\ln p_S$ and $\ln S$ are collinear. In this case, I already showed in section 4.3 that the remaining parameters are not identified; the Hessian of the likelihood equation is singular. When $\gamma > 0$ however, the model does not reduce to the linear utility

model and the Hessian of the likelihood equation is non-singular. Hence, the scaled hedonic location choice parameters $\widehat{\eta}_2$, $\widehat{\eta}_3$, and γ are identified. The computational results in the next section illustrate this fact.

Next consider estimation of the education production function. Since $\ln S$ is endogenous, I need to use the location choice equation to produce a valid instrument for $\ln S$. In the linear utility model, this is not possible since the system of empirical equations is linear and both equations contain the same set of exogenous variables. Here however, the set of equations is not linear. $\ln S$, the consumer's choice of location, depends on x_1 and x_2 in a non-linear way. The set of equations can be estimated either jointly or by creating an instrument for $\ln S$ by projecting $\ln S$ on high order polynomials in x_1 and x_2. By construction, this instrument is independent of z_2 and is not a linear function of x_1 and x_2.

Thus, the non-linearity in the econometric system induced by the equilibrium restrictions solves both the hedonic estimation problem and the endogenous regressor estimation problem. Moreover, this non-linearity is not arbitrary but is produced by the structure of the model. Different structures that result in different non-linearities of this type can be tested against the data and can be used to derive valid inferences about the structural param-

eters defining the economy.

5.5 Empirical results

To test the identification results from the previous section, I simulated datasets in economies in which $\gamma > 0$, and then estimated the structural parameters using maximum likelihood. The likelihood equation depends on the solution to the pricing equation. In order to make this calculation computationally feasible, I computed the solution to the pricing equation using a projection method.[19] Details of the method used are given in appendix E.

Some results are given in Tables 6-9. Tables 6 and 8 show results when I assume that $\ln s_1$, the child's educational outcome is observed. Tables 7 and 9 show results when I assume that the data is censored so that I observe $\ln s_1$ only if $s_1 \leq 12$. Otherwise I observe $(s_1 \geq 12)$. These tables are included to show how the model performs when the available data are censored.

All parameter estimates are within two standard deviations of the true parameter values. The method successfully estimates the structural parameters of the model using synthetic data.

Table 10 displays preliminary results from an analysis of the National Ed-

[19] See Judd (1998).

ucational Longitudinal Study. These preliminary results show that although the method produces parameter estimates, it is likely that the specification used in this paper is not general enough to match the data. As seen in Table 10, the parameter estimates are implausible. The elasticity of children's educational attainment with respect to school quality is much too large. Two generalizations of the specification developed in this paper that could account for these implausible estimates are that the distribution of consumer traits are not normally distributed or that the education production function is not a constant elasticity function. Nesheim (2002) is pursuing these generalizations to test whether these richer specifications better match the data.

6 Conclusion and future work

I developed a theoretical model in which heterogeneous consumers sort into locations based on the average education of the residents in those locations. The model provides a rich theoretical and empirical basis to analyze peoples' location choices. I find sufficient conditions for an equilibrium to exist in this model and analyze some of the patterns that result in an equilibrium of this

model.

I examined two versions of this model to gain an understanding of their theoretical and empirical properties. Both of these examples shed light on the issues that affect identification in more general hedonic estimation problems and in neighborhood effects estimation problems. In the restricted model in which utility is linear in consumption, I find that the model is not identified and discuss how this identification result is related to the hedonic estimation literature. I estimate the model in this case and trace out the subspaces of the parameter space that are consistent with the data. Few restrictions on the underlying parameters result.

In the more general version, I find that the model is identified. In particular, knowledge of the non-linear functional form of the equilibrium price function is used to show that the system is identified. This result illustrates how specification of the structure of a hedonic economy and imposition of the equilibrium restrictions implied by such an economy can be used to achieve identification of the structural parameters describing the economy. I test the model with synthetic and real data and find that the estimation procedure produces reliable estimates of the underlying structural parameters.

Ongoing work in Nesheim (2002) extends this model to test two kinds of

specification errors. First, it tests alternate functional forms for the education production and the utility function to see how robust the results are to changes in the functional forms tested. Second, it tests alternate distributions of the consumer traits to see how robust the results are to assumptions about these distributions. All of these tests are implemented with the theoretical and computational methods developed in this dissertation.

Other future work is needed to analyze more complicated models in which consumers' choice of location depends not only on a one-dimensional index of neighborhood quality but rather on a multi-dimensional index. This work will require further development of the methods developed in this dissertation.

A Proofs

Proof of Theorem 3.1. The proof is by construction. Let μ_1 and σ_{11} be the mean and variance of x_1 and let $f(z)$ be the solution to the following differential equation:

$$\begin{aligned} f'(z) &= \frac{h(z)}{\phi(\mu_1 + f, \mu_1, \sigma_{11}) + \phi(\mu_1 - f, \mu_1, \sigma_{11})} \\ f(0) &= 0 \end{aligned}$$

Since the right side is positive and satisfies a Lipschitz condition, the differential equation has a unique, strictly increasing global solution. Let $F(x) = f^{-1}\left(\left|x_1 - \mu_1 - \frac{1}{2}\sigma_{11}\right|\right)$ and let $q(z) = 0$ for all z. Then $E\left[e^{x_1} \left| x \in F^{-1}(z)\right.\right] = \overline{S}$ for all z. Moreover, conditions 1, 3, and 4 in definition 3.1 are trivially satisfied, and condition 2 is satisfied because $h(z) = dP(z)$ almost everywhere.

■

Proof of Lemma 3.2. I must prove that condition 1 of definition 3.1 holds for $\left(\widehat{F}, \widehat{q}\right)$. Define

$$\mathcal{S} = \left\{S \left| S = E\left[e^{x_1} \left| x \in F^{-1}(z) \text{ for some } z\right.\right]\right.\right\}$$

and define $\widehat{\mathcal{S}}$ analogously for $\widehat{F}$. First, I show that $\mathcal{S} = \widehat{\mathcal{S}}$. If $S_1 \in \mathcal{S}$, then by definition $S_1 = E\left[e^{x_1} | x \in F^{-1}(z_1)\right]$ for some z_1. Therefore, $F^{-1}(z_1) \neq \emptyset$. Let $x_1 \in F^{-1}(z_1)$ and let $z_2 = \widehat{F}(x_1)$, then by assumption 1 $S_1 = E\left[e^{x_1} | x \in \widehat{F}^{-1}(z_2)\right]$ and so $S_1 \in \widehat{\mathcal{S}}$. Now suppose $S_1 \in \widehat{\mathcal{S}}$. By definition, $S_1 = E\left[e^{x_1} | x \in \widehat{F}^{-1}(z_1)\right]$ for some z_1 and $\widehat{F}^{-1}(z_1) \neq \emptyset$. Let $x_1 \in \widehat{F}^{-1}(z_1)$ and suppose $S_1 \notin \mathcal{S}$. Let $z_2 = F(x_1)$, then $S_1 \neq S_2 = E\left[e^{x_1} | x \in F^{-1}(z_2)\right]$. But by assumption 1), this implies that $S_2 = E\left[e^{x_1} | x \in \widehat{F}^{-1}(z_1)\right]$ contradicting the supposition that $S_1 = E\left[e^{x_1} | x \in \widehat{F}^{-1}(z_1)\right]$.

Now define

$$p(S) = \left\{ \; q(z) \quad \text{if } S = E\left[e^{x_1} | x \in F^{-1}(z)\right] \; \right\}$$

and define $\widehat{p}$ analogously for $\left(\widehat{F}, \widehat{q}\right)$. Assumption 2 guarantees that $p(S) = \widehat{p}(S)$ for all $S \in \mathcal{S}$.

Finally, by assumption 1 if equilibrium (F, q) assigns a consumer x to a location with quality S_1, then $\left(\widehat{F}, \widehat{q}\right)$ also assigns the consumer to a location with quality S_1. Since $(S_1, p(S_1))$ was optimal when the consumer faced the set of choices $\mathcal{S}$ with price schedule p, it remains optimal under $\left(\widehat{F}, \widehat{q}\right)$ since $\mathcal{S} = \widehat{\mathcal{S}}$ and $p = \widehat{p}$. Thus, condition 1 in definition 3.1 is satisfied and $\left(\widehat{F}, \widehat{q}\right)$

is a locational equilibrium. ■

Proof of Theorem 3.3. Let $\mathcal{S}_1 = \{S \,|\mu_d(S) > 0\}$. Then $\mathcal{S}_1$ is a countable set and $\mu_d(\mathcal{S}_1) \leq 1$. Assuming $\mathcal{S}_1$ is non-empty, define $z_{-1} = 0$ and for each $S_i \in \mathcal{S}_1$ define $z_i(S_i)$ recursively so that

$$\int_{z_{i-1}}^{z_i} h(z)\, dz = \mu_d(S_i)$$

Then, on the domain $[z_{i-1}, z_i]$ define $f_i(z)$ to be the solution to

$$\begin{aligned} f_i'(z) &= \frac{h(z)}{\frac{\phi(\ln(S_i+f_i(z)))}{S_i+f_i(z)} + \frac{\phi(\ln(S_i-f_i(z)))}{S_i-f_i(z)}} \\ f_i(z_{i-1}) &= 0 \end{aligned}$$

Since the right side is positive and satisfies a Lipschitz condition, the differential equation has a unique, strictly increasing global solution. For $x \in G^{-1}(S_i)$, let $F(x) = f_i^{-1}(|e^{x_1} - S_i|)$ and let $q(z) = p(S_i)$ for all $z \in [z_{i-1}, z_i]$.

Of course, if $\mathcal{S}_1$ is empty, the above construction is superfluous.

Next let $d\mu_a(S)$ represent the density of μ_a and consider

$$\mathcal{S}_2 = \{S \,|d\mu_a(S) > 0\}$$

$\mathcal{S}_2$ can be represented by a countable union of connected intervals with non-overlapping interiors. Let $\left[S_{j-1}^a, S_j^a\right]$ be the $j'th$ such interval, let $z_{-1}^a = \max_i \{z_i\}$ and let

$$\int_{z_{j-1}^a}^{z_j^a} h(z) = \int_{S_{j-1}^a}^{S_j^a} d\mu_a(S)\, dS$$

On the domain, $\left[z_{j-1}^a, z_j^a\right]$ let $g_j(z)$ solve

$$\begin{aligned} g_j'(z) &= \frac{h(z)}{d\mu_a(g_j(z))} \\ g_j\left(z_{j-1}^a\right) &= S_{j-1}^a \end{aligned}$$

For all $x \in G^{-1}\left(\left[S_{j-1}^a, S_j^a\right]\right)$, define $F(x) = g_j^{-1}(G(x))$ and for $z \in \left[z_{j-1}^a, z_j^a\right]$ define $q(z) = p(g_j(z))$.

By construction, (F, q) is a locational equilibrium and satisfies conditions 1 and 2 given in the statement of the theorem. ■

Proof of Theorem 5.1. Lemmas D.1 and D.2 in appendix D imply that $\lim_{f\to-\infty} F(f) = 0$ and $\lim_{f\to\infty} F(f) = \infty$. Since $\frac{dF(f)}{df} > 0$ and since $F \in C^2(R)$, the implicit function theorem then implies that there exists a unique $\widehat{f}(S) \in C^2(R_{++})$ satisfying $S - F(f) = 0$ where $F(f)$ is defined in (5.7)

such that $\frac{d\widehat{f}(S)}{dS} > 0$. Furthermore, this implies that $\lim_{S\to 0}\widehat{f}(S) = -\infty$ and $\lim_{S\to\infty}\widehat{f}(S) = \infty$. As a result, the differential equation

$$\begin{aligned}\widehat{f}(S) &= (1-\eta_1)\ln S + \gamma p + \ln p_S - \ln(A_0\eta_1)\\ p(0) &= 0\end{aligned}$$

has a unique solution $\widehat{p} \in C^2(R_{++}) \bigcap C^0(R_+)$. By construction $\widehat{p}$ satisfies the equilibrium condition (5.6) for all $S \geq 0$. Moreover, given this price function, every consumers' maximization problem has a unique solution. To see this substitute $\widehat{f}(S)$ into the first-order condition (5.2) to obtain the equivalent condition

$$-e^{-\gamma(e^{x_2}-\widehat{p})}\widehat{p}_S\left(1 - h(x)\, e^{\gamma e^{x_2} - \widehat{f}(S)}\right) = 0$$

where $h(x) = e^{\eta_2 x_1 + x_3 + x_4}$. Since $\widehat{p}_S > 0$, $\lim_{S\to 0}\widehat{f}(S) = -\infty$, $\lim_{S\to\infty}\widehat{f}(S) = \infty$, and $\frac{d\widehat{f}(S)}{dS} > 0$, the equation has a unique solution $S^*(x)$ for every consumer x. In addition, marginal utility is positive for all $S < S^*(x)$ and marginal utility is negative for all $S < S^*(x)$. ■

Proof of Theorem 5.2. I will construct a discontinuous function $\widehat{f}(S)$ that in turn defines a price function that satisfies the equilibrium conditions

given in problem 5.1.

Since $L_1 > 0$, lemma D.1 implies that$\frac{dF(f)}{df} > 0$ when $f << 0$. Furthermore lemma D.2 implies that $\frac{dF(f)}{df} > 0$ when $f >> 0$. Therefore, $\frac{dF(f)}{df}$ has a least two zeros. Assume for the moment that there are only two zeros, f_1 and f_2 where $f_1 < f_2$. Then $\frac{dF(f)}{df} > 0$ for all $f \in (-\infty, f_1) \bigcup (f_2, \infty)$. By the implicit function theorem, there is a unique function $\widehat{f_1}(S)$ that satisfies $S = F(f)$ for all $S < F(f_1)$. This function satisfies $\widehat{f_1}(S) < f_1$ for all $S < F(f_1)$ and $\widehat{f_1}(F(f_1)) = f_1$. Moreover, $\frac{d\widehat{f_1}(S)}{dS} > 0$ for all $S < F(f_1)$. Similarly, there is a unique function $\widehat{f_2}(S)$ that satisfies $S = F(f)$ for all $S > F(f_2)$. This function satisfies $\widehat{f_2}(S) > f_2$ for all $S > F(f_2)$ and $\widehat{f_2}(F(f_2)) = f_2$. Also, $\frac{d\widehat{f_2}}{dS} > 0$ for all $S > F(f_2)$. Since $F(f_2) < F(f_1)$, the equation $S = F(f)$ has multiple solutions for every $S \in [F(f_2), F(f_1)]$.

Let $\widehat{f}(S)$ be a function of the form

$$\widehat{f}(S) = \left\{ \begin{array}{ll} \widehat{f_1}(S) & S < S_0 \\ \widehat{f_2}(S) & S > S_0 \end{array} \right\} \tag{A.1}$$

where $S_0 \in [F(f_2), F(f_1)]$. By construction, such a function satisfies the pricing condition $S - F(f) = 0$ for all $S \neq S_0$; where $F(f)$ is given in (5.7). At the point, S_0, however, this pricing condition does not characterize

the equilibrium because $\widehat{f}(S)$ jumps from $\widehat{f}_1(S_0)$ to $\widehat{f}_2(S_0)$ where $\widehat{f}_1(S_0) < \widehat{f}_2(S_0)$. Looking back at the consumer first-order condition (5.2), the function $\widehat{f}(S)$ is an equilibrium if

$$F\left(\widehat{f}_1(S_0)\right) = E\left[e^{x_1} \left| g_0\left(\widehat{f}_1(S_0), x\right) \leq 0 \leq g_0\left(\widehat{f}_2(S_0), x\right)\right.\right] \tag{A.2}$$

The left side of $(A.2)$ is a continuous function of the single variable S_0 defined on the compact domain $[F(f_2), F(f_1)]$. Let $S_1 = F(f_1)$. For all $f \in \left(\widehat{f}_1(S_1), \widehat{f}_2(S_1)\right)$,

$$E\left[e^{x_1} | g_0(f, x) = 0\right] = F(f) < F\left(\widehat{f}_1(S_1)\right)$$

implying that

$$F\left(\widehat{f}_1(S_1)\right) > E\left[e^{x_1} \left| g_0\left(\widehat{f}_1(S_1), x\right) \leq 0 \leq g_0\left(\widehat{f}_2(S_1), x\right)\right.\right]$$

On the other hand, let $S_2 = F(f_2)$. For all $f \in \left(\widehat{f}_1(S_2), \widehat{f}_2(S_2)\right)$,

$$E\left[e^{x_1} | g_0(f, x) = 0\right] = F(f) > F\left(\widehat{f}_2(S_2)\right)$$

implying that

$$F\left(\widehat{f}_2\left(S_2\right)\right) < E\left[e^{x_1}\left|g_0\left(\widehat{f}_1\left(S_2\right), x\right) \leq 0 \leq g_0\left(\widehat{f}_2\left(S_2\right), x\right)\right.\right]$$

Thus, by the intermediate value theorem, there must be an $S_0 \in [S_2, S_1]$ that exactly satisfies $(A.2)$.

By construction, the function $\widehat{f}(S)$ defined by $(A.1)$ with discontinuity point S_0 defined by $(A.2)$ satisfies the conditions for an equilibrium. The associated price function is $\widehat{p}(S)$ where

$$\widehat{p}(S) = \left\{ \begin{array}{ll} \widehat{p}_1(S) & \text{if } S \leq S_0 \\ \widehat{p}_2(S) & \text{if } S \geq S_0 \end{array} \right\}$$

and $\widehat{p}_1(S)$ solves

$$\widehat{f}_1(S) = (1-\eta_1)\ln S + \gamma p + \ln p_S$$

$$p(0) = 0$$

for all $S \leq S_0$ and $\widehat{p}_2(S)$ solves

$$\widehat{f_2}(S) = (1-\eta_1)\ln S + \gamma p + \ln p_S$$

$$p(S_0) = \widehat{p}_1(S_0)$$

for all $S \geq S_0$.

If $\frac{dF(f)}{df}$ has more than two zeros, the same arguments can be repeated possibly resulting in a larger number of discontinuity points. ■

B Derivation of the equilibrium ODE in the exponential utility model

As in section 5.1, let

$$f = \gamma p + \ln p_S + (1-\eta_1)\ln S - \ln(A_0\eta_1)$$

be an index of neighborhood quality and price. The equilibrium differential equation is

$$S = E\left[e^{x_1} \,|\, g_0(f,x) = 0\right] \tag{B.1}$$

where $g_0(f,x) = f - \eta_2 x_1 - \gamma e^{x_2} - x_3 - x_4$. This section derives a formula for the right side of $(B.1)$.

First, define the inverse transformation

$$\begin{aligned} x_1 &= z_1 \\ x_2 &= z_2 \\ x_3 &= z_3 \\ x_4 &= z_4 + f - \eta_2 z_1 - \gamma e^{z_2} - z_3 \end{aligned}$$

with Jacobian

$$\left| \frac{\partial(x_1, x_2, x_3, x_4)}{\partial(z_1, z_2, z_3, z_4)} \right| = \begin{vmatrix} 1 & 0 & 0 & -\eta_2 \\ 0 & 1 & 0 & -\gamma e^{z_2} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1$$

Then, since $x \sim N(\mu, \Sigma)$, z has the probability density function

$$\psi(z_1, z_2, z_3, z_4) = \phi_4(z_1, z_2, z_3, z_4 + f - \eta_2 z_1 - \gamma e^{z_2} - z_3, \nu, \Omega)$$

where $\phi_4(\cdot, \nu, \Omega)$ is the four-dimensional normal probability density function with mean ν and variance Ω and ν and Ω are the mean and variance of the first four components of the vector x.

In terms of the variable z, the right side of the differential equation $(B.1)$ is equivalent to

$$E\left[e^{z_1} \,|z_4 = 0\right] \tag{B.2}$$

which by using the definition of conditional density can be written

$$E\left[e^{z_1}|z_4 = 0\right] = \frac{\iiint e^{z_1}\phi_4(z_1, z_2, z_3, g, \nu, \Omega)\, dz_1 dz_2 dz_3}{\iiint \phi_3(z_1, z_2, z_3, g, \nu, \Omega)\, dz_1 dz_2 dz_3} \tag{B.3}$$

where $g = f - \eta_2 z_1 - \gamma e^{z_2} - z_3$. Factoring the normal density, equation $(B.3)$ becomes

$$E\left[e^{z_1}|z_4 = 0\right] = \frac{\int \phi(z_2, \nu_2, \omega_{22}) \left(\iint e^{z_1}\phi_3(z_1, z_3, g, \xi, \Lambda)\, dz_1 dz_3\right) dz_2}{\int \phi(z_2, \nu_2, \omega_{22}) \left(\iint \phi_2(z_1, z_3, g, \xi, \Lambda)\, dz_1 dz_3\right) dz_2} \tag{B.4}$$

where

$$\xi = \left(\nu_1 + \frac{\omega_{12}}{\omega_{22}}(z_2 - \nu_2), \nu_3 + \frac{\omega_{23}}{\omega_{22}}(z_2 - \nu_2), \nu_4 + \frac{\omega_{24}}{\omega_{22}}(z_2 - \nu_2)\right)'$$

and

$$\Lambda = \begin{bmatrix} \omega_{11} - \frac{\omega_{12}^2}{\omega_{22}} & \omega_{13} - \frac{\omega_{12}\omega_{23}}{\omega_{22}} & \omega_{14} - \frac{\omega_{14}\omega_{24}}{\omega_{22}} \\ \omega_{13} - \frac{\omega_{12}\omega_{23}}{\omega_{22}} & \omega_{33} - \frac{\omega_{23}^2}{\omega_{22}} & \omega_{34} - \frac{\omega_{23}\omega_{24}}{\omega_{22}} \\ \omega_{14} - \frac{\omega_{14}\omega_{24}}{\omega_{22}} & \omega_{34} - \frac{\omega_{23}\omega_{24}}{\omega_{22}} & \omega_{44} - \frac{\omega_{24}^2}{\omega_{22}} \end{bmatrix}$$

Let $I_1(z_2)$ represent the inner pair of integrals in the numerator of $(B.4)$. $I_1(z_2)$ can be rewritten as

$$I_1(z_2) = \iint\limits_{z_1,z_3} \frac{e^{z_1 - 0.5qf'Q'\Lambda^{-1}Qq}}{(2\pi)^{1.5}|\Lambda|^{0.5}} dz_1 dz_3 \tag{B.5}$$

where $q = (z_1 - \xi_1, z_3 - \xi_3, f - \eta_2\xi_1 - \gamma e^{z_2} - \xi_3 - \xi_4)'$ and

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\eta_2 & -1 & 1 \end{bmatrix}$$

Letting $T = Q^{-1}\Lambda(Q')^{-1}$ and factoring the integrand, this integral can be

further simplified to

$$I_1(z_2) = \frac{e^{-0.5\frac{q_3^2}{\tau_{33}}}}{\sqrt{2\pi\tau_{33}}} \iint\limits_{z_1,z_3} e^{z_1}\phi_2(z_1, z_3, \pi, \chi)\, dz_1 dz_3 \tag{B.6}$$

where $\pi = \left(\xi_1 + \frac{\tau_{13}}{\tau_{33}}q_3, \xi_3 + \frac{\tau_{13}}{\tau_{33}}q_3\right)'$ and $\chi = \begin{bmatrix} \tau_{11} - \frac{\tau_{13}^2}{\tau_{33}} & \tau_{12} - \frac{\tau_{13}\tau_{23}}{\tau_{33}} \\ \tau_{12} - \frac{\tau_{13}\tau_{23}}{\tau_{33}} & \tau_{22} - \frac{\tau_{23}^2}{\tau_{33}} \end{bmatrix}$.
This integral reduces to the analytic expression

$$I_1(z_2) = \frac{e^{-0.5\frac{q_3^2}{\tau_{33}} + \pi_1 + 0.5\chi_{11}}}{\sqrt{2\pi\tau_{33}}} \tag{B.7}$$

Thus, the conditional expectation in equation $(B.3)$ reduces to the ratio of two one-dimensional integrals

$$E\left[e^{z_1}|z_4 = 0\right] = \frac{\int \frac{e^{-\frac{1}{2}\frac{(z_2-\nu_2)^2}{\omega_{22}} + \pi_1 + \frac{1}{2}\chi_{11} - \frac{1}{2}\frac{q_3^2}{\tau_{33}}}}{2\pi\sqrt{\omega_{22}\tau_{33}}} dz_2}{\int \frac{e^{-\frac{1}{2}\frac{(z_2-\nu_2)^2}{\omega_{22}} - \frac{1}{2}\frac{q_3^2}{\tau_{33}}}}{2\pi\sqrt{\omega_{22}\tau_{33}}} dz_2} \tag{B.8}$$

or

$$E\left[e^{z_1}|z_4 = 0\right] = \frac{e^{\nu_1 + \frac{1}{2}\omega_{11}} \int e^{-\frac{1}{2}\frac{(z_2-\nu_2-\omega_{12})^2}{\omega_{22}} - \frac{1}{2}\frac{(q_3-\tau_{13})^2}{\tau_{33}}} dz_2}{\int e^{-\frac{1}{2}\frac{(z_2-\nu_2)^2}{\omega_{22}} - \frac{1}{2}\frac{q_3^2}{\tau_{33}}} dz_2} \tag{B.9}$$

where

$$q_3 = f - \eta_2\xi_1 - \xi_3 - \xi_4 - \gamma e^{z_2}$$

$$f = \gamma p + \ln p_S + (1-\eta_1)\ln S - \ln(A_0\eta_1)$$

$$\xi_1 = \nu_1 + \tfrac{\omega_{12}}{\omega_{22}}(z_2 - \nu_2)$$

$$\xi_3 = \nu_3 + \tfrac{\omega_{23}}{\omega_{22}}(z_2 - \nu_2)$$

$$\xi_4 = \nu_4 + \tfrac{\omega_{24}}{\omega_{22}}(z_2 - \nu_2)$$

$$\tau_{13} = \eta_2\Lambda_{11} + \Lambda_{12} + \Lambda_{13}$$

$$\tau_{33} = \eta_2^2\Lambda_{11} + 2\eta_2(\Lambda_{12} + \Lambda_{13}) + \Lambda_{22} + \Lambda_{33} + 2\Lambda_{23}$$

Thus, the differential equation $(B.1)$ can be written as

$$S - F(f(S, p, p_S)) = 0 \tag{B.10}$$

where

$$F(f) = \frac{e^{\nu_1 + \frac{1}{2}\omega_{11}} \int e^{-\frac{1}{2}\frac{(z_2-\nu_2-\omega_{12})^2}{\omega_{22}} - \frac{1}{2}\frac{(q_3(f)-\tau_{13})^2}{\tau_{33}}} dz_2}{\int e^{-\frac{1}{2}\frac{(z_2-\nu_2)^2}{\omega_{22}} - \frac{1}{2}\frac{q_3(f)^2}{\tau_{33}}} dz_2} \tag{B.11}$$

and where $q_3(f) = f - \eta_2\xi_1 - \xi_3 - \xi_4 - \gamma e^{z_2}$ and

$$f(S, p, p_S) = (1-\eta_1)\ln S + \gamma p + \ln p_S - \ln(A_0\eta_1)$$

C Computation of the integrand

Consider integration of the integral appearing in the numerator in (5.7) and $(B.11)$. This integral is

$$I(f) = \int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{(z_2-\nu_2-\omega_{12})^2}{\omega_{22}}-\frac{1}{2}\frac{(q_3(f)-\tau_{13})^2}{\tau_{33}}} dz_2 \tag{C.1}$$

Since the integrand resembles the kernel of a Gaussian density and since the range of integration is doubly infinite, the most obvious transformation of variables is $z_2 = \sqrt{2}\sigma(f) z_3 + \nu(f)$ where $\nu(f) = \arg\max_{z_2}\{J(z_2, f)\}$,

$$J(z_2, f) = -\frac{1}{2}\frac{(z_2-\nu_2-\omega_{12})^2}{\omega_{22}} - \frac{1}{2}\frac{(q_3(f)-\tau_{13})^2}{\tau_{33}} \tag{C.2}$$

and $\sigma(f) = \sqrt{\frac{d^2 J(\nu,f)}{dz_2^2}}$. Using this transformation the integral can be rewritten as

$$I(f) = \int_{-\infty}^{\infty} e^{-z_3^2} \cdot \left(\sqrt{2}\sigma(f) e^{z_3^2 + J\left(\sqrt{2}\sigma(f) z_3 + \nu(f), f\right)}\right) dz_3 \tag{C.3}$$

This transformation allows use of Gauss-Hermite integration techniques, centers the bulk of the computational effort around the peak of the integrand, and seeks to transform the integrand into a less steeply peaked integrand.

This transformation did not perform well in practice since the integrand often has two peaks. So I tried another transformation which performed much better. Consider the transformation $z_2 = c_1(f) \ln\left(\frac{1+z_3}{1-z_3}\right) + c_2(f)$ where c_1 and c_2 are chosen conservatively so that $(\nu(f) - 2\sigma(f), \nu(f) + 2\sigma(f))$ is mapped into $(-0.5, 0.5)$.[20] This transformation flattens the integrand, maps most of the mass of the integrand into the range $(-0.5, 0, 5)$ and proved successful in practice. The Jacobian of the transformation is $\frac{2c_1(f)}{1-z_3^2}$ and the integral $I(f)$ in $(C.1)$ becomes

$$I(f) = \int_{-1}^{1} \frac{2c_2(f)}{\sqrt{1-z_3^2}} \frac{e^{J(z_2,f)}}{\sqrt{1-z_3^2}} dz_3 \tag{C.4}$$

A similar transformation applies to the denominator. With these transformations equation (5.7) can be approximated quickly and accurately. Moreover, the calculated integral is a smooth function of both the parameters and f.

[20] An alternative transformation is $z_2 = \frac{c_1 z_3}{\sqrt{1-z_3^2}} + c_2$ where $c_1 > 0$.

D Laplacian approximation to $F(f)$ when $|f| \rightarrow \infty$

The function to be approximated is

$$F(f) = \frac{N(f)}{D(f)} \tag{D.1}$$

where

$$N(f) = e^{\nu_1 + \frac{1}{2}\omega_{11}} \int_{z_2} e^{-\frac{1}{2}\frac{(z_2 - \nu_2 - \omega_{12})^2}{\omega_{22}} - \frac{1}{2}\frac{(q_2(f) - \tau_{12})^2}{\tau_{22}}} dz_2$$

$$D(f) = \int_{z_2} e^{-\frac{1}{2}\frac{(z_2 - \nu_2)^2}{\omega_{22}} - \frac{1}{2}\frac{q_2(f)^2}{\tau_{22}}} dz_2$$

and

$q_2(f) = \frac{f - \widehat{\eta}_2\left(\nu_1 - \frac{\omega_{12}}{\omega_{22}}\nu_2\right) - k_1 z_2 - \gamma e^{z_2}}{\sqrt{\omega_{33}}}$

$f = (1 - \eta_1)\ln S + \gamma p + \ln p_S - \widehat{\eta}_0$

$\widehat{\eta}_0 = \frac{1}{2}\omega_{55} + \ln \eta_1 + \alpha_1 + \beta_1$

$k_1 = \widehat{\eta}_3 + \widehat{\eta}_2 \frac{\omega_{12}}{\omega_{22}}$

$\widehat{\eta}_2 = \eta_2 + \alpha_2 + \beta_2$

$\widehat{\eta}_3 = \alpha_3 + \beta_3$

$\tau_{12} = \frac{\widehat{\eta}_2}{\sqrt{\omega_{33}}}\Lambda_{11}$

$\tau_{22} = \frac{\widehat{\eta}_2^2}{\omega_{33}}\Lambda_{11} + 1$

$\Lambda_{11} = \omega_{11} - \frac{\omega_{12}^2}{\omega_{22}}$

Throughout this section I assume that $\widehat{\eta_2} > 0$ and $k_1 > 0$. These conditions are economically plausible since they imply that both parental education and income have a net positive effect on children's schooling.

When $|f| \to \infty$ both $N(f) \to 0$ and $D(f) \to 0$. Therefore, a direct approximation of $F(f)$ is impossible and a Laplace type approximation is required. I develop this approximation in this section.

The numerator $N(f)$ in $(D.1)$can be rewritten as

$$N(f) = L_N \int_{-\infty}^{\infty} e^{-g^N(f,z_2)} dz_2 \tag{D.2}$$

where

$L_N = e^{\nu_1 + \frac{1}{2}\omega_{11}}$

$g^N(f, z_2) = \frac{(z_2 - n_0)^2}{2\omega_{22}} + \frac{\left(n_1(f) - \frac{k_1}{\sqrt{\omega_{33}}} z_2 - \frac{\gamma}{\sqrt{\omega_{33}}} e^{z_2}\right)^2}{2\tau_{22}}$

$n_0 = \nu_2 + \omega_{12}$

$n_1(f) = \frac{f - \widehat{\eta}_2\left(\nu_1 - \frac{\omega_{12}}{\omega_{22}}\nu_2\right)}{\sqrt{\omega_{33}}} - \tau_{12}$

$k_1 = \widehat{\eta}_3 + \widehat{\eta}_2 \frac{\omega_{12}}{\omega_{22}}$

In the limit as $f \to -\infty$, $n_1(f)$ approaches $-\infty$. Moreover, the minimizer of $g^N(f, z_2)$ approaches $-\infty$, $g^N(f, z_2)$ is nearly quadratic near this minimizer, and the minimizer is increasing in f. Away from its minimum, $g^N(f, z_2)$ grows at least as fast as z_2^2 and so $e^{-g^N(f,z_2)}$ shrinks at least as quickly as $e^{-z_2^2}$. Therefore, we can approximate the integral when f is small with a Laplace type approximation. A Taylor approximation of g^N around its minimizer is

$$\begin{aligned} \widetilde{g}^N(f, z_2) &= g_0^N(f) + \frac{1}{2} g_2^N(f) \cdot (z_2 - z_N)^2 + \\ &\quad \frac{1}{6} g_3^N(f) \cdot (z_2 - z_N)^3 + \frac{1}{24} g_4^N(f) \cdot (z_2 - z_N)^4 \end{aligned} \tag{D.3}$$

where z_N is the minimizer of $g^N(f, z_2)$ and $g_i^N(f)$ is the $i'th$ derivative of $g^N(f, z_2)$ with respect to z_2 evaluated at z_N. Using this Taylor approximation, letting $z_3 = z_2 - z_N$, and further taking a sixth-order Taylor approximation to $e^{-\frac{1}{6} g_3^N(f) z_3^3 - \frac{1}{24} g_4^N(f) z_3^4}$, an approximation to the integral in (D.2)

is

$$\widetilde{N}(f) = L_N e^{-g_0^N(f)} \int_{-\infty}^{\infty} e^{-g_2^N(f)\frac{z_3^2}{2}} \begin{pmatrix} 1 - g_3^N(f)\frac{z_3^3}{6} - \\ g_4^N(f)\frac{z_3^4}{24} + \left(g_3^N(f)\right)^2 \frac{z_3^6}{72} \end{pmatrix} dz_3 \quad \text{(D.4)}$$

Using the transformation $z = \sqrt{\frac{g_2^N(f)}{2}} z_3$ this is equivalent to

$$\widetilde{N}(f) = \frac{\sqrt{2} L_N e^{-g_0^N(f)}}{\sqrt{g_2^N(f)}} \int_{-\infty}^{\infty} e^{-z^2} \begin{pmatrix} 1 - \\ g_3^N(f)\left(\frac{2}{g_2^N(f)}\right)^{1.5} \frac{z^3}{6} - \\ g_4^N(f)\left(\frac{2}{g_2^N(f)}\right)^2 \frac{z^4}{24} + \\ \left(g_3^N(f)\right)^2 \left(\frac{2}{g_2^N(f)}\right)^3 \frac{z^6}{72} \end{pmatrix} dz \quad \text{(D.5)}$$

which reduces to

$$\widetilde{N}(f) = \frac{\sqrt{2\pi} L_N e^{-g_0^N(f)}}{\sqrt{g_2^N(f)}} \left(1 - \frac{g_4^N(f)}{8\left(g_2^N(f)\right)^2} + \frac{5}{24} \frac{\left(g_3^N(f)\right)^2}{\left(g_2^N(f)\right)^3}\right) \quad \text{(D.6)}$$

Formulas for the functions $g_i^N(f)$ are available from the author upon request.

The denominator $D(f)$ in $(D.1)$ is approximated in a similar way. The

denominator is

$$D(f) = \int_{-\infty}^{\infty} e^{-g^D(f,z_2)} dz_2 \tag{D.7}$$

where

$g^D(f, z_2) = \frac{(z_2 - d_0)^2}{2\omega_{22}} + \frac{\left(d_1(f) - \frac{k_1}{\sqrt{\omega_{33}}} z_2 - \frac{\gamma}{\sqrt{\omega_{33}}} e^{z_2}\right)^2}{2\tau_{22}}$

$d_0 = \nu_2$

$d_1(f) = \frac{f - \widehat{\eta}_2\left(\nu_1 - \frac{\omega_{12}}{\omega_{22}}\nu_2\right)}{\sqrt{\omega_{33}}}$

$k_1 = \widehat{\eta}_3 + \widehat{\eta}_2 \frac{\omega_{12}}{\omega_{22}}$

I form the Taylor approximation

$$\begin{aligned} \widetilde{g}^D(f, z_2) &= g_0^D(f) + \frac{1}{2} g_2^D(f) \cdot (z_2 - z_D)^2 + \\ &\quad \frac{1}{6} g_3^D(f) \cdot (z_2 - z_D)^3 + \frac{1}{24} g_4^D(f) \cdot (z_2 - z_D)^4 \end{aligned} \tag{D.8}$$

where z_D is the minimizer of $g^D(f, z_2)$ and $g_i^D(f)$ is the $i'th$ derivative of $g^D(f, z_2)$ with respect to z_2 evaluated at z_D. Then letting $z_3 = z_2 - z_{D,}$ an approximation to the integral in $(D.7)$ is

$$\widetilde{D}(f) = e^{-g_0^D(f)} \int_{-\infty}^{\infty} e^{-g_2^D(f)\frac{z_3^2}{2}} \begin{pmatrix} 1 - g_3^D(f)\frac{z_3^3}{6} - \\ g_4^D(f)\frac{z_3^4}{24} + g_3^D(f)^2 \frac{z_3^6}{72} \end{pmatrix} dz_3 \tag{D.9}$$

Making the change of variable $z = \sqrt{\frac{g_2^D(f)}{2}} z_3$, this is equivalent to

$$\widetilde{D}(f) = \frac{\sqrt{2} e^{-g_0^D(f)}}{\sqrt{g_2^D(f)}} \int_{-\infty}^{\infty} e^{-z^2} \left(\begin{array}{c} 1- \\ \left(\frac{2}{g_2^D(f)}\right)^{\frac{3}{2}} g_3^D(f) \frac{z^3}{6} - \\ \left(\frac{2}{g_2^D(f)}\right)^{2} g_4^D(f) \frac{z^4}{24} + \\ \left(\frac{2}{g_2^D(f)}\right)^{3} g_3^D(f)^2 \frac{z^6}{72} \end{array} \right) dz \tag{D.10}$$

which reduces to

$$\widetilde{D}(f) = \frac{\sqrt{2\pi} e^{-g_0^D(f)}}{\sqrt{g_2^D(f)}} \left(1 - \frac{g_4^D(f)}{8 g_2^D(f)^2} + \frac{5}{24} \frac{g_3^D(f)^2}{g_2^D(f)^3} \right) \tag{D.11}$$

Formulas for the functions $g_i^D(f)$ are available from the author upon request.

Using the approximations $\widetilde{N}(f)$ and $\widetilde{D}(f)$, the approximation to the function $F(f)$ is

$$\widetilde{F}(f) = \frac{\frac{L_N e^{-g_0^N(f)}}{\sqrt{g_2^N(f)}} \left(1 - \frac{g_4^N(f)}{8 g_2^N(f)^2} + \frac{5}{24} \frac{g_3^N(f)^2}{g_2^N(f)^3} \right)}{\frac{e^{-g_0^D(f)}}{\sqrt{g_2^D(f)}} \left(1 - \frac{g_4^D(f)}{8 g_2^D(f)^2} + \frac{5}{24} \frac{g_3^D(f)^2}{g_2^D(f)^3} \right)} \tag{D.12}$$

The limit of this approximation is characterized by the following two lemmas.

Lemma D.1 *Let $F(f)$ be as given in equation $(D.1)$. Then*

$$\begin{aligned} \lim_{f\to-\infty} F(f) &= \\ \lim_{f\to-\infty}\left[L_0 e^{\nu_1+\frac{1}{2}\omega_{11}+L_1 f}\right] &= 0 \end{aligned}$$

where

$$L_0 = e^{-\frac{1}{2}\frac{\left(\tau_{12}+\frac{k_1}{\sqrt{\omega_{33}}}\omega_{12}\right)^2}{\tau_{22}+\frac{k_1^2}{\omega_{22}}\omega_{22}} - -\frac{\left(\tau_{12}+\frac{k_1}{\sqrt{\omega_{33}}}\omega_{12}\right)\frac{(\hat{\eta}_2\nu_1+\hat{\eta}_3\nu_2)}{\sqrt{\omega_{33}}}}{\tau_{22}+\frac{k_1^2}{\omega_{33}}\omega_{22}}}$$

$$L_1 = \frac{\tau_{12}+\frac{k_1}{\sqrt{\omega_{33}}}\omega_{12}}{\tau_{22}+\frac{k_1^2}{\omega_{33}}\omega_{22}}\left(\frac{1}{\sqrt{\omega_{33}}}\right)$$

Proof. By construction $F(f)$ can be made arbitrarily close to the function $\widetilde{F}(f)$ given in equation $(D.12)$ by choosing f small enough. Moreover, when $f \to -\infty$ then $d_1(f) \to -\infty$. Analysis of $(D.12)$ and the functions $g_i^D(f)$ and $g_i^N(f)$ (formulas available upon request) shows that in the limit $\widetilde{F}(f)$ reduces to

$$\widetilde{F}(f) = L_N e^{g_0^D(f)-g_0^N(f)} \tag{D.13}$$

where

$$g_0^D(f) = \frac{(z_D(f) - d_0)^2}{2\omega_{22}} + \frac{\left(d_1(f) - \frac{k_1}{\sqrt{\omega_{33}}} z_D(f) - \frac{\gamma}{\sqrt{\omega_{33}}} e^{z_D(f)}\right)^2}{2\tau_{22}}$$

$$g_0^N(f) = \frac{(z_N(f) - n_0)^2}{2\omega_{22}} + \frac{\left(n_1(f) - \frac{k_1}{\sqrt{\omega_{33}}} z_N(f) - \frac{\gamma}{\sqrt{\omega_{33}}} e^{z_N(f)}\right)^2}{2\tau_{22}}$$

and

$$d_0 = \nu_2$$

$$d_1(f) = \frac{f - \widehat{\eta}_2\left(\nu_1 - \frac{\omega_{12}}{\omega_{22}}\nu_2\right)}{\sqrt{\omega_{33}}}$$

$$n_0 = d_0 + \omega_{12}$$

$$n_1(f) = d_1 - \tau_{12}$$

$$z_D(f) = \frac{d_0\tau_{22} + \frac{k_1}{\sqrt{\omega_{33}}}\omega_{22} d_1(f)}{\tau_{22} + \frac{k_1^2}{\omega_{33}}\omega_{22}}$$

$$z_N(f) = \frac{n_0\tau_{22} + \frac{k_1}{\sqrt{\omega_{33}}}\omega_{22} n_1(f)}{\tau_{22} + \frac{k_1^2}{\omega_{33}}\omega_{22}}$$

Evaluating $g_0^D(f) - g_0^N(f)$ gives

$$\widetilde{F}(f) = L_0 e^{L_1 f} \tag{D.14}$$

where

$$L_0 = e^{\nu_1+\frac{1}{2}\omega_{11}-\frac{1}{2}\frac{\left(\tau_{12}+\frac{k_1}{\sqrt{\omega_{33}}}\omega_{12}\right)^2}{\tau_{22}+\frac{k_1^2}{\omega_{22}}\omega_{22}}-\frac{\left(\tau_{12}+\frac{k_1}{\sqrt{\omega_{33}}}\omega_{12}\right)\frac{(\widehat{\eta}_2\nu_1+\widehat{\eta}_3\nu_2)}{\sqrt{\omega_{33}}}}{\tau_{22}+\frac{k_1^2}{\omega_{33}}\omega_{22}}}$$

$$L_1 = \frac{\tau_{12}+\frac{k_1}{\sqrt{\omega_{33}}}\omega_{12}}{\tau_{22}+\frac{k_1^2}{\omega_{33}}\omega_{22}}\left(\frac{1}{\sqrt{\omega_{33}}}\right)$$

Since I assume $\widehat{\eta}_2 > 0$ and $k_1 > 0$, this implies that $\lim_{f\to-\infty}\widetilde{F}(f) = 0$. ■

Lemma D.2 $\lim_{f\to\infty} F(f) = \infty$ *and* $\lim_{f\to\infty}\frac{dF(f)}{df} > 0$.

Proof. By construction $\lim_{f\to\infty} F(f) = \widetilde{F}(f)$ where $\widetilde{F}(f)$ is given in equation $(D.12)$. Also, by definition of $d_1(f)$ when $f \to \infty$, $d_1(f) \to \infty$. Analysis of $\widetilde{F}(f)$ and the functions $g_i^D(f)$ and $g_i^N(f)$ (formulas available upon request) shows that in the limit $\widetilde{F}(f)$ reduces to

$$\widetilde{F}(f) = L_N e^{g_0^D(f)-g_0^N(f)} \tag{D.15}$$

where

$$g_0^D(f) = \frac{(z_D-d_0)^2}{2\omega_{22}} + \frac{\left(d_1-\frac{k_1}{\sqrt{\omega_{33}}}z_D-\frac{\gamma}{\sqrt{\omega_{33}}}e^{z_D}\right)^2}{2\tau_{22}}$$

$$g_0^N = \frac{(z_N - n_0)^2}{2\omega_{22}} + \frac{\left(n_1 - \frac{k_1}{\sqrt{\omega_{33}}} z_N - \frac{\gamma}{\sqrt{\omega_{33}}} e^{z_N}\right)^2}{2\tau_{22}}$$

and

$$d_0 = \nu_2$$

$$d_1(f) = \frac{f - \widehat{\eta}_2\left(\nu_1 - \frac{\omega_{12}}{\omega_{22}}\nu_2\right)}{\sqrt{\omega_{33}}}$$

$$n_0 = d_0 + \omega_{12}$$

$$n_1(f) = d_1 - \tau_{12}$$

Let $z_D^0 = \ln\left(\frac{\sqrt{\omega_{33}} d_1}{\gamma}\right)$ and let $z_N^0 = \ln\left(\frac{\sqrt{\omega_{33}} n_1}{\gamma}\right)$ then when f is large

$$g_0^D = \frac{(z_D - d_0)^2}{2\omega_{22}} + \frac{\left(d_1 z_D^0 - \left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1\right) z_D\right)^2}{2\tau_{22}} \tag{D.16}$$

$$g_0^N = \frac{(z_N - n_0)^2}{2\omega_{22}} + \frac{\left(n_1 z_N^0 - \left(\frac{k_1}{\sqrt{\omega_{33}}} + n_1\right) z_N\right)^2}{2\tau_{22}}$$

and therefore

$$\begin{aligned} z_D &\cong \frac{\tau_{22}d_0 + \omega_{22}z_D^0 d_1\left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1\right)}{\tau_{22} + \omega_{22}\left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1\right)^2} \\ z_N &\cong \frac{\tau_{22}n_0 + \omega_{22}z_N^0 n_1\left(\frac{k_1}{\sqrt{\omega_{33}}} + n_1\right)}{\tau_{22} + \omega_{22}\left(\frac{k_1}{\sqrt{\omega_{33}}} + n_1\right)^2} \end{aligned}$$

As a result,

$$\begin{aligned} g_0^D &= \frac{1}{2}\frac{\left(d_1 z_D^0 - \left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1\right)d_0\right)^2}{\tau_{22} + \omega_{22}\left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1\right)^2} \\ g_0^N &= \frac{1}{2}\frac{\left(n_1 z_N^0 - \left(\frac{k_1}{\sqrt{\omega_{33}}} + n_1\right)n_0\right)^2}{\tau_{22} + \omega_{22}\left(\frac{k_1}{\sqrt{\omega_{33}}} + n_1\right)^2} \end{aligned} \tag{D.17}$$

and

$$\begin{aligned} g_0^D - g_0^N &= \frac{1}{2}\frac{\left(d_1 z_D^0 - \left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1\right)d_0\right)^2}{\tau_{22} + \omega_{22}\left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1\right)^2} - \\ &\quad \frac{1}{2}\frac{\left(\begin{array}{c}(d_1 - \tau_{12})\, z_N^0 - \\ \left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1 - \tau_{12}\right)(d_0 + \omega_{12})\end{array}\right)^2}{\tau_{22} + \omega_{22}\left(\frac{k_1}{\sqrt{\omega_{33}}} + d_1 - \tau_{12}\right)^2} \end{aligned} \tag{D.18}$$

Letting $b = k_1 + d_1$, this is equivalent to

$$g_0^D - g_0^N = \frac{1}{2}\frac{\left(d_1 z_D^0 - b d_0\right)^2}{\tau_{22} + \omega_{22} b^2} - \frac{1}{2}\frac{\left(\begin{array}{c}\left(d_1 - \tau_{12}\right) z_N^0 - \\ \\ \left(b - \tau_{12}\right)\left(d_0 + \omega_{12}\right)\end{array}\right)^2}{\tau_{22} + \omega_{22}\left(b - \tau_{12}\right)^2} \tag{D.19}$$

which when expanded becomes a fourteen term expression. Twelve of the fourteen terms equal zero. The remaining two are non-zero and the final expression is

$$\begin{aligned} \lim_{f \to \infty} \left[g_0^D(f) - g_0^N(f)\right] &= \\ \lim_{f \to \infty} \left[\frac{\omega_{12} z_N^0(f)}{\omega_{22}} - \frac{1}{2}\frac{\left(2 d_0 \omega_{12} + \omega_{12}^2\right)}{\omega_{22}}\right] &= \infty \end{aligned} \tag{D.20}$$

■

E Computing approximate solutions to $f(S)$

The function to be approximated $f(S, \theta)$ solves

$$S - F(f, \theta) = 0$$

defined on the domain $S \in R_+$. The equation $F(f)$ is defined in (5.7). For a slow pointwise approximation to this equation one can use any nonlinear equation solver. However, for empirical purposes a fast smooth functional approximation is desirable. To calculate this approximation, I approximated $f(S)$ by a piecewise polynomial function and used Chebyshev collocation to calculate the coefficients of the approximating polynomial.

I first truncated the domain so that $S \in [S_L, S_H]$. Since the approximation is used for empirical work no values of $f(S)$ are needed for $S < S_L$ and $S > S_H$. Then I decomposed this domain into several (typically 5) subdomains. I summarize the approximation for a typical subdomain. Let its endpoints be S_A and S_B and let S' be a vector whose M components are the M zeros of the derivative of the degree $M+1$ Chebyshev polynomial adapted to the interval $[S_A, S_B]$. In particular, $S'(1) = S_A$ and $S'(M) = S_B$. Let $T(S'(i))$ be the vector whose row m is the degree m Chebyshev polynomial

evaluated at the point $S'(i)$. Let a_1 be a vector of real numbers of length m. Define the approximation

$$\widehat{f}(S'(i)) = a_1' T(S'(i)) \tag{E.1}$$

The coefficients of the approximation are then computed by finding the vector a_1 that solves the system of m nonlinear equations

$$\begin{aligned} f_0 &= a_1' T(S_A) \\ f_1 &= a_1' T(S_B) \\ S'(i) &= F(a_1' T(S'(i))) \text{ for } i = 2, ..., m-1 \end{aligned} \tag{E.2}$$

References

[1] Bartik, Timothy (1987), "Estimating Hedonic Demand Parameters with Single Market Data: The Problems Caused by Unobserved Tastes," *The Review of Economics and Statistics, 69(1):178-180.*

[2] Betts, Julian (1995), "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth," *Review of Economics and Statistics*, 77:231-250.

[3] Brock, William, and Durlauf, Steven (2000), "Interactions-Based Models," forthcoming in *Handbook of Econometrics 5.*

[4] Brown, James, and Rosen, Harvey (1982), "On the Estimation of Structural Hedonic Price Models," *Econometrica*, 50(3):765-768.

[5] Card, David, Krueger, Alan (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100(1):1-40.

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

[7] Epple, Dennis and Platt, Glenn (1998), "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in Preferences and Incomes," *Journal of Urban Economics*, 43(1):23-51.

[8] Epple, Dennis, and Sieg, Holger (1999a), "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, 99(4):828-858.

[9] Epple, Dennis, and Sieg, Holger (1999b), "The Tiebout Hypothesis and Majority Rule: An Empirical Analysis," working paper.

[10] Hanushek, Eric (1996), "Measuring Investment in Education," *Journal of Economic Perspectives*, 10(4):9-30.

[11] Haveman, Robert, and Wolfe, Barbara (1995), "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature, 33(4):1829-78.*

[12] Heckman, James, Layne-Farrar, Anne, and Todd, Petra (1996), "Does measureed School Quality Really Matter? An Examination of the Earnings-Quality Relationship," in *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, ed. Burtless, Gary, Washington, DC, Brookings Institution, 192-289.

[13] Ekeland, Ivar, Heckman, James, and Nesheim, Lars (2001), "Theory of Identification in Hedonic Equilibrium Models," unpublished working paper.

[14] Jencks, Christopher, and Mayer, Susan (1990), "Social Consequences of Growing Up in a Poor Neighborhood," in *Inner-City Poverty in the United States*, eds. Jencks, Christopher, and Peterson, Paul, Washington, D.C.: Brookings Institution.

[15] Judd, Kenneth (1998), *Numerical Methods in Economics*, Cambridge, MA, MIT Press.

[16] Kahn, Shulamit, and Lang, Kevin (1988), "Efficient Estimation of Structural Hedonic Systems," *International Economic Review*, 29:157-166.

[17] Kniesner, Thomas, and Leeth, John (1995), *Simulating Workplace Safety Policy*, Boston and Dordrecht: Kluwer Academic Publishing.

[18] Nesheim, Lars (2002), "Empirical Estimation of School Quality Effects in a Locational Choice Equilibrium Model," unpublished working paper.

[19] Rosen, Sherwin (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82:34-50.

[20] Sattinger, Michael (1980), *Capital and the Distribution of Labor Earnings*, Amsterdam: North-Holland.

[21] Teulings, Coen (1995), "The Wage Distribution in a Model of the Assignment of Skills to Jobs," *Journal of Political Economy*, 103(2):280-315.

[22] Tinbergen, Jan (1959), "On the Theory of Income Distribution," in *Selected Papers of Jan Tinbergen*, edited by L. H. Klaassen, L. M. Koyck, and H. J. Witteveen, Amsterdam: North-Holland.

F Figures

Figure 1: Equilibrium price function

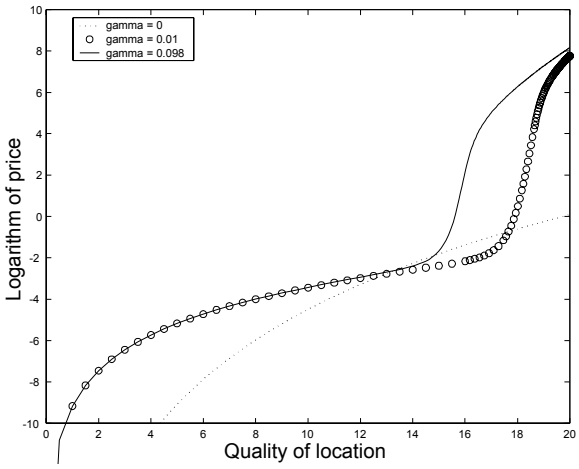

Figure 2: Slope of equilibrium price function

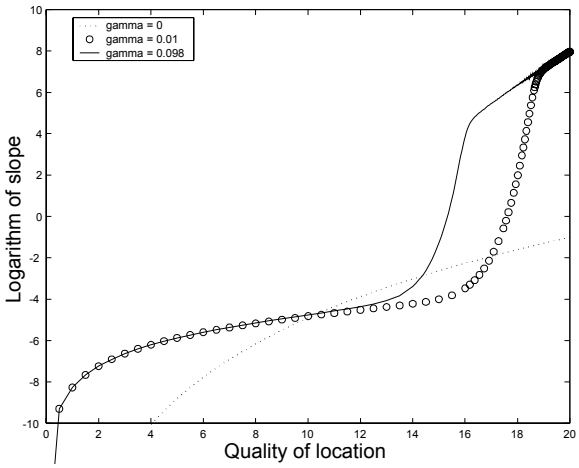

Figure 3: Conditional mean of log-income

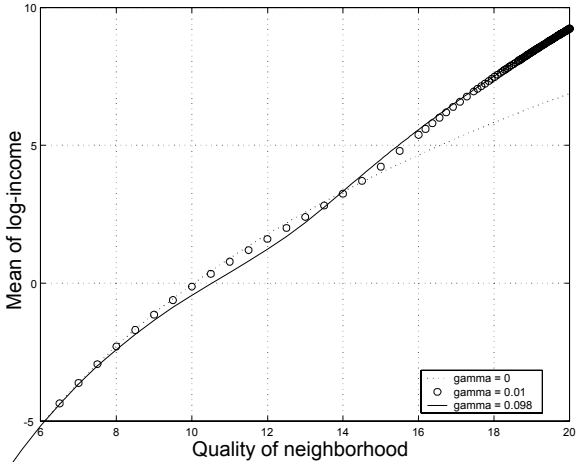

Figure 4: Conditional mean of z_2

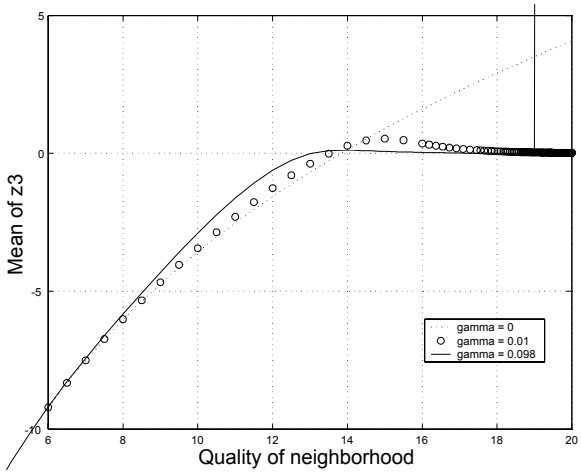

Figure 5: Conditional variance of log-schooling

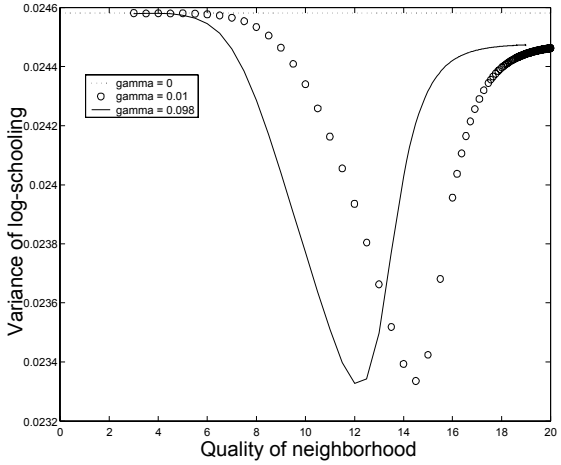

Figure 6: Conditional variance of log-income

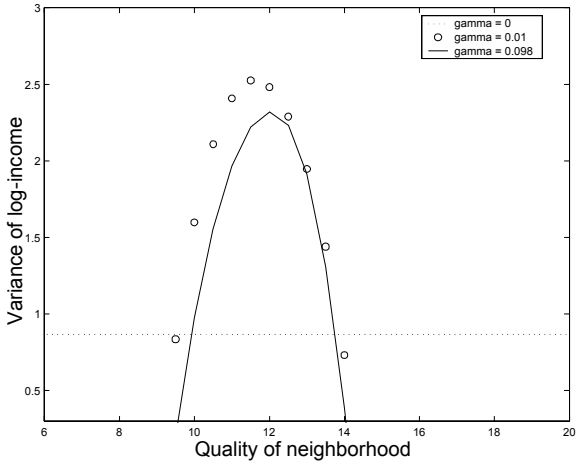

Figure 7: Conditional variance of z_2