TECHNICAL WORKING PAPER SERIES

NON-PARAMETRIC DEMAND ANALYSIS WITH AN APPLICATION TO THE DEMAND FOR FISH

Joshua D. Angrist
Kathryn Graddy
Guido W. Imbens

Technical Working Paper No. 178

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1995

Data used in this paper are available on request (e-mail requests to econkg@vax.ox.ac.uk). Thanks go to Gary Chamberlain for a number of helpful discussions and comments and to seminar participants at the Harvard-MIT Econometrics workshop and Princeton University for helpful comments and suggestions. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Technical Working Paper #178
April 1995

NON-PARAMETRIC DEMAND ANALYSIS WITH AN APPLICATION TO THE DEMAND FOR FISH

ABSTRACT

Instrumental variables (IV) estimation of a demand equation using time series data is shown to produce a weighted average derivative of heterogeneous potential demand functions. This result adapts recent work on the causal interpretation of two-stage least squares estimates to the simultaneous equations context and generalizes earlier research on average derivative estimation to models with endogenous regressors. The paper also shows how to compute the weights underlying IV estimates of average derivatives in a simultaneous equations model. These ideas are illustrated using data from the Fulton Fish market in New York City to estimate an average elasticity of wholesale demand for fresh fish. The weighting function underlying IV estimates of the demand equation is graphed and interpreted. The empirical example illustrates the essentially local and context-specific nature of instrumental variables estimates of structural parameters in simultaneous equations models.

Joshua D. Angrist
Department of Economics
Hebrew University
Jerusalem 91905
ISRAEL
and NBER

Kathryn Graddy
Department of Economics
Jesus College
Oxford University
Oxford OX1 3DW
UNITED KINGDOM

Guido W. Imbens
Department of Economics
Harvard University
Cambridge, MA 02138
and NBER

1. Introduction

Simultaneous Equations Models (SEM) are often used to characterize causal relationships in econometrics. In spite of the central role that the SEM plays in theoretical discussions of econometric modelling, as an empirical tool the model appears to have important limitations. First and most importantly, one must have valid instruments. Second, identification in the SEM appears to be tied to linearity of the structural equations and to the assumption of constant coefficients. For example, the textbook rank and order conditions are defined for linear models with constant coefficients.

Linearity of all relationships in the SEM is clearly not essential for identification. Kelejian (1971) pointed out that linearity of reduced-form conditional expectations is not required for identification of structural equations, nor does it render conventional two-stage least squares (2SLS) estimates inconsistent. But the analysis of non-linear structural equations is more complicated and the non-linear SEM is almost always presented with fixed coefficients and additive error terms in the structural equation. For example, Amemiya's non-linear 2SLS estimator (Amemiya 1974) is designed for a model of this type and most of the recent work on non-parametric simultaneous equations models imposes an additive error structure.[1]

Linear models with constant coefficients are more convenient to estimate than non-linear models, and non-linear models with additive errors are more tractable than a general non-linear model. Applied researchers often view these restricted models as approximations to richer models allowing for non-linear and heterogeneous response functions. We show that in principle this interpretation is correct, with the caveat that for a given set of behavioral relationships, the

---

[1]Examples include Hausman and Newey, 1995; Newey, Powell, and Vella, 1994; Newey and Powell, 1989; and Roehrig, 1988.

nature of the approximation varies with the choice of instrument. In particular, we show that additive errors and a parametric functional form are not necessary for conventional IV or 2SLS estimators to identify interesting aspects of the structural relationships of interest in a general non-linear SEM. For example, given a supply shift that is excluded from the demand equation, conventional instrumental variables techniques identify a weighted average of the derivative of a time-varying heterogeneous demand function. This generalizes recent results on average derivative estimation (Stoker 1986; Powell, Stock, and Stoker 1989) to models with endogenous regressors.

There are two components to the average derivatives discussed here. First, there is averaging along the length of a nonlinear response function. Second, there is averaging of the function derivative as the function shifts from period to period (or market to market). An important feature of our approach is the weighting function underlying IV estimates can be computed and used to describe which part of a heterogeneous and time-varying demand function is being captured by the estimates.

The ideas in the paper are illustrated using data from the Fulton Fish market in New York City to estimate the average elasticity of demand for fresh fish sold in this wholesale market. These data were first analyzed by Graddy (1995). This example is useful for our purposes for two reasons. First, weather conditions at sea appear to provide valid and powerful instrumental variables that shift the supply curve for fresh fish. Second, Graddy's analysis suggests that white and Asian buyers at the Fulton market have different demand functions and enter the market in different proportions from day to day. The market demand function is therefore likely to exhibit time-varying heterogeneity generated from a changing customer mix.

2. Demand and Supply Functions

We begin by summarizing a framework for nonparametric demand analysis. The particular relationship of interest is the aggregate demand for fish at the Fulton fish market. The theoretical demand and supply relationships consist of functions showing what the quantity supplied or the quantity demanded would be for any price, including (but not only) the observed price. These relationships are denoted by

$$q_t^d(p, z, x) \tag{1}$$

$$q_t^s(p, z, x); \text{ for } t = 1, \ldots, T.$$

The subscript t can be thought of as indexing distinct markets in a cross-section or indexing days, as in our time series application. The scalar covariates z and x are candidate instruments because they shift the supply and demand curves.

System (1) describes what quantity demanded and quantity supplied would have been in period t if consumers and producers had been confronted with prices p, and shift variables z and x. Since only $q_t^d(p_t, z_t, x_t)$ is actually observed and it is not possible to change history, we say that $q_t^d(p, z_t, x_t)$ is *potential* quantity demanded at $p \neq p_t$. This representation of the demand and supply functions generalizes the usual linear SEM, in which potential quantities are linear and parallel functions of price. A typical linear system is

$$q_t^d(p, z, x) = \alpha_0 + \alpha_1 p + \alpha_2 z + \alpha_3 x + \epsilon_t$$

$$q_t^s(p, z, x) = \beta_0 + \beta_1 p + \beta_2 z + \beta_3 x + \eta_t,$$

where ϵ_t and η_t are scalar error terms assumed to be independent of z_t and x_t. Here shifts in the demand and supply curves are solely attributable to the additive error terms, as illustrated in Figure (1a) for a demand curve without exogenous covariates.

A more general version of the SEM allows for non-parallel shifts in demand and supply, while maintaining linearity for changes in quantity demanded and quantity supplied:

$$q_t^d(p, z, x) = \alpha_0 + \alpha_{1t}p + \alpha_{2t}z + \alpha_{3t}x + \epsilon_t$$

$$q_t^s(p, z, x) = \beta_0 + \beta_{1t}p + \beta_{2t}z + \beta_{3t}x + \eta_t.$$

Two demand curves corresponding to this case are illustrated in Figure (1b). Such non-parallel shifts could be a consequence of aggregation if, for example, the market demand elasticity depends on the ethnic mix of buyers each day (see, e.g., Stoker 1993).

The nonlinear generalization of the linear SEM, as analyzed by Amemiya (1974) and Newey and Powell (1989), is

$$q_t^d(p, z, x) = q^d(p, z, x) + \epsilon_t$$

$$q_t^s(p, z, x) = q^s(p, z, x) + \eta_t.$$

As in the linear SEM, in this case shifts in demand and supply are parallel, as illustrated in Figure (1c). Roehrig (1988) discusses sufficient conditions for this model to be nonparametrically identified. Note that additivity of the error terms means that the average (over periods) demand function is non-stochastic and equal to $q^d(p, z, x)$.[2]

Finally, allowing for non-additive error terms gives:

$$q_t^d(p, z, x) = q^d(p, z, x, \epsilon_t), \tag{2a}$$

$$q_t^s(p, z, x) = q^s(p, z, x, \eta_t). \tag{2b}$$

System (2) allows for both non-linear changes in quantity demanded and non-parallel shifts in demand, as illustrated in Figure (1d). This error-term formulation can approximate system (1)

[2]Use of additive errors in an otherwise non-linear model is a not an innocuous restriction because the functions $q^d(p, z, x)$ and $q^s(p, z, x)$ are not conditional expectations.

arbitrarily well because the error terms can be thought of as containing information about changing function shapes as well as random shifts in position.

Note that the concept of an average slope has two components in system (1) or (2). First, there is the average slope along the length of a non-linear demand curve within periods. Second, there is the average slope across demand curves at a given price.

2.1 The interpretation of potential demand

Rubin (1990) traces the notion of potential outcomes in statistics to Jerzy Neyman's discussion of "potential yield" in agricultural experiments. Similarly, we view potential quantities demanded and supplied as quantities that could have been revealed by experimental manipulation of prices on day t. The interpretation of simultaneous equation models as describing potential or counterfactual realizations that could be revealed by experimentation appears to originate with Haavelmo (1944):

"When we set up a system of theoretical relationships and use economic names for the otherwise purely theoretical variables involved, we have in mind some actual *experiment*, or some *design of an experiment*, which we could at least imagine arranging, in order to measure those quantities in real economic life that we think might obey the laws imposed on their theoretical namesakes. (Haavelmo, 1944, page 6; italics in original.)

In economic terms, such an experiment would generate movement along a given demand or supply curve.[3] Note that Haavelmo's interpretation of simultaneous equations models does not require the relationships of interest to be either linear or parallel. Even in linear models, time-varying composition effects could occur as a consequence of aggregation (Stoker 1993.)

[3]For more on this point, see Heckman (1992, p. 879-880), who interprets Haavelmo's discussion of demand equations as "defining and determining counterfactual states of the economy."

2.2 Instrumental variables in a non-parametric model

We begin with two preliminary conditions:

Assumption 1. For all (p,z,x),

a. $\{[q_t^d(p, z, x), q_t^s(p, z, x), p_t, z_t, x_t]; t = 1, \ldots, T\}$ is a stationary ergodic sequence with finite first and second moments.

b. The functions $q_t^d(p, z, x)$, $q_t^s(p, z, x)$ are continuously differentiable in p.

The purpose of Assumption 1a is to facilitate inference and to be precise when using stochastic concepts to discuss identification and inference. Assumption 1b guarantees that the concept of an average derivative is well-defined.

Initially, z_t is taken to be a binary instrument with positive probability of being either zero or one. The conditions that makes z_t an instrument for the effect of price on quantity demanded are:

Assumption 2.

a. (independence) For all values in the support of (p, z, x), z_t is jointly independent of the pair of functions $\{q_t^d(p, z, x_t), q_t^s(p, z, x_t)\}$ given x_t.

b. (exclusion) For all p,

$$q_t^d(p, 1, x_t) = q_t^d(p, 0, x_t) \equiv q_t^d(p, x_t).$$

c. For some periods,

$$q_t^s(p_t, 1, x_t) \neq q_t^s(p_t, 0, x_t).$$

Assumption 2a is equivalent to independence of the instruments and structural errors in the standard linear SEM. Casting the independence assumption in terms of potential outcomes in this manner is a version of the ignorability concept discussed by Rubin (1978). Ignorability of instrumental variables is discussed in Angrist, Imbens and Rubin (1995).

To estimate the elasticity of demand for fish at the Fulton fish market, we use weather conditions at sea to construct instruments for prices. The first part of Assumption 2 requires that these weather conditions be "as good as randomly assigned," conditional on covariates. In other words, weather conditions should not be related to the quantity of fish that would be demanded or supplied at any p,z combination given x_t. As a consequence, comparisons of average prices on stormy and clear days and comparisons of average quantities on stormy and clear days provide an unbiased estimate of the causal effect of weather conditions on prices and quantities.

Assumption 2b requires that the instrument be unrelated to quantity demanded through any mechanism other than prices. In the linear SEM, a similar condition is typically imposed through zero restrictions on structural coefficients or through the rank and order conditions on the reduced form. In our application, the exclusion restriction requires weather conditions at sea to be unrelated to dealers' willingness to buy fish at a given price, or to the number of dealers transacting at the Fulton market. This is a strong assumption that could be violated if, for example, weather conditions at sea are correlated with weather conditions on shore and shore weather affects market demand. Another problem is that buyers might try to use knowledge of weather conditions to obtain market power. For example, if buyers know that the weather has been good for fishing, they could try to collude and obtain a lower price.

Assumption 2c requires the instrument to have an effect on quantity supplied, at least some of the time. The instrument will then affect prices through the reduced form. Maintaining the first part of Assumption 2 (independence), it is straightforward to test whether weather conditions affect prices and quantities.

3. Identification

We focus on the demand equation and begin by considering the identifying power of a binary instrument. Suppose that $z_t = 1$ to indicate stormy weather conditions at sea and $z_t = 0$ otherwise, and suppose that Assumption 2 holds with x_t equal to a constant. Given this modification and the exclusion restriction, the demand equation can be written $q^d_t(p)$ and the supply equation can be written $q^s_t(p, z)$. The observed price, p_t, is assumed to clear the market so that

$$q^d_t(p_t) - q^s_t(p_t, z_t) = 0.$$

This equilibrium condition defines the reduced form for a binary instrument as follows:

Assumption R1. For all t, there are unique numbers p_{0t} and p_{1t} such that:

$$q^d_t(p_{0t}) - q^s_t(p_{0t}, 0) = 0 \tag{3}$$

$$q^d_t(p_{1t}) - q^s_t(p_{1t}, 1) = 0.$$

This implies that it is always possible to find a unique equilibrium price given $z_t = 0$ or $z_t = 1$.[4] We can express observed prices in terms of potential prices as follows:

[4]Sufficient conditions for this are: upward sloping supply, downward sloping demand, and continuous differentiability of demand and supply functions in prices.

$$p_t = h_t(z_t) \equiv p_{0t} + (p_{1t} - p_{0t})z_t. \qquad (4)$$

The potential prices, p_{0t} and p_{1t}, play a key role in our interpretation of simple IV estimators. Note that p_{0t} and p_{1t} are random variables independent of z_t, even though the observed price, $p_t = p_{0t} + (p_{1t} - p_{0t})z_t$ is clearly not independent of z_t. Also, only one of the potential prices is actually observed in a given period.

3.1 The Wald Estimator

The IV estimator using z_t and a constant as an instrument for a regression of quantities on prices is the ratio of the covariance of (z_t, q_t) to the covariance of (z_t, p_t). This is the Wald estimator based on comparisons of observed quantities and prices under different weather conditions:

$$\hat{\alpha}_{1,0} = \frac{[\Sigma_t z_t q_t/\Sigma_t z_t] - [\Sigma_t(1-z_t)q_t/\Sigma_t(1-z_t)]}{[\Sigma_t z_t p_t/\Sigma_t z_t] - [\Sigma_t(1-z_t)p_t/\Sigma_t(1-z_t)]}\ .$$

Given Assumption 1, we can apply the ergodic theorem (e.g., White 1984, page 42) to show that $\hat{\alpha}_{1,0}$ converges (as $T \to \infty$) almost surely to $\alpha_{1,0}$ where

$$\alpha_{1,0} = \frac{E[q_t|\ z_t=1] - E[q_t|\ z_t=0]}{E[p_t|\ z_t=1] - E[p_t|\ z_t=0]}\ . \qquad (5)$$

Our first identification results concern the nature of $\alpha_{1,0}$. Provided the instrument potentially leads either to an increase in prices in all periods or to a decrease in prices in all periods, $\alpha_{1,0}$ can be shown to be a particular weighted average of the derivative of the demand function, $\partial q_t^d(p)/\partial p$.[5] The idea that a binary instrument has a monotone impact on prices is captured by the following assumption:

Assumption 3 (monotonicity for a binary instrument.)

Either $p_{1t} \geq p_{0t}$ holds for all t or $p_{1t} \leq p_{0t}$ for all t.

Without loss of generality, we assume here that $p_{1t} \geq p_{0t}$. This means that stormy fishing weather in any period leads to an equilibrium price at least as high as would have been observed if the weather had been clear in the same period. Although this condition has some testable implications in combination with Assumption 2, it cannot be checked directly because only one price is observed each period. On the other hand, we show in the next section that Assumption 3 is plausible and economically meaningful in a market model such as the one analyzed here.[6]

The following lemma shows that comparisons of expected quantities at different values of z_t are determined solely by the derivative of the demand function over the range (p_{0t}, p_{1t}):

Lemma 1. $E[q_t | z_t=1] - E[q_t | z_t=0] = E\{ \int_{p_{0t}}^{p_{1t}} [\partial q_t^d(s)/\partial s] \, ds \}$. (6)

Proof. The demand function for period t can be written:

---

[5]Even if $q_t^d(p)$ is truly linear, say $q_t^d(p) = \alpha_t + \beta_t p$, the weighting scheme matters because β_t may differ across periods.

[6]Imbens and Angrist (1994), Angrist and Imbens (1995), and Angrist, Imbens and Rubin (1995) discuss monotonicity in the context of models for program evaluation. Manski (1995) discusses the use of an ordered-outcomes assumption to bound simultaneous-equations bias in a general non-linear model.

$$q_t^d(p_t) = q_t^d(0) + \int_0^{p_t} [\partial q_t^d(s)/\partial s]ds$$

$$= q_t^d(0) + \{ \int_0^{p_{1t}} [\partial q_t^d(s)/\partial s]ds \}z_t + \{ \int_0^{p_{0t}} [\partial q_t^d(s)/\partial s]ds \}(1-z_t)$$

By virtue of Assumption 2a (independence), the random variables $(q_t^d(0), [\partial q_t^d(s)/\partial s], p_{1t}, p_{0t})$ are jointly independent of the realized value of z_t. Therefore,

$$E[q_t | z_t=1] - E[q_t | z_t=0] =$$

$$E \{ \int_0^{p_{1t}} [\partial q_t^d(s)/\partial s]ds - \int_0^{p_{0t}} [\partial q_t^d(s)/\partial s]ds \}.$$

Using Assumption 3, which requires $p_{1t} \geq p_{0t}$, completes the proof. □

Lemma 1 means that regardless of the marginal distribution of prices, comparisons of average quantities conditional on z_t are informative about the slope of the demand curve only over the range of price variation induced by the instrument. $\alpha_{1,0}$, normalizes this comparison by the difference in average prices conditional on the instrument. The following proposition shows that for periods where prices are affected by the instrument, this normalized comparison is a conditional weighted average of the demand curve slope over the range $[p_{0t}, p_{1t}]$:

Proposition 1. Let $F_1(p)$ be the CDF of p_{1t} and let $F_0(p)$ be the CDF of p_{0t}. These CDFs are time-invariant by virtue of Assumption 1 and the definition of p_{1t} and p_{0t}. Then,

$$\alpha_{1,0} = \int_0^{\infty} E[\partial q_t^d(s)/\partial s | p_{1t} \geq s > p_{0t}]\omega(s)\, ds, \tag{7}$$

where

$$\omega(s) = [F_0(s) - F_1(s)]/[E(p_t | z_t=1) - E(p_t | z_t=0)],$$

with $\int_0^\infty \omega(s)ds = 1$ and $\omega(s) \geq 0$ for all s.

Proof. The proposition is established by evaluating the right hand side of (6). Note that

$$E\{\int_{p_{0i}}^{p_{1i}} \partial q_i^d(s)/\partial s\, ds\} = E\{\int_0^\infty [\partial q_i^d(s)/\partial s][1(p_{1i} \geq s > p_{0i})]\, ds\},$$

where $1(p_{1i} \geq s > p_{0i})$ is an indicator for p_{0i} less than s and p_{1i} at least s.

Taking expectations of $[\partial q_i^d(s)/\partial s][1(p_{1i} \geq s > p_{0i})]$ inside the integral gives

$$E[\partial q_i^d(s)/\partial s \mid p_{1i} \geq s > p_{0i}][F_0(s) - F_1(s)],$$

since, under monotonicity, $Pr[1(p_{1i} \geq s > p_{0i})=1] = F_0(s) - F_1(s)$. Note also that monotonicity implies the numerator and denominator of $\omega(s)$ are both positive. It remains to show that $\omega(s)$ integrates to 1. Note that the denominator is:

$$E[p_i \mid z_i=1] - E[p_i \mid z_i=0] = E[p_{1i}] - E[p_{0i}].$$

Using the fact that the mean of a positive random variable is the integral of 1 minus the CDF, we have

$$E[p_{1i}] - E[p_{0i}] = \int_0^\infty [F_0(s) - F_1(s)]ds.$$

□

Proposition 1 characterizes the two types of averaging implicit in IV estimation. First, the IV estimate reflects a conditional average slope at each price, $E[\partial q_i^d(s)/\partial s \mid p_{1i} \geq s > p_{0i}]$. This average is computed for fixed s, using the conditional distribution of $\partial q_i^d(s)/\partial s$ given $p_{1i} \geq s > p_{0i}$ as the demand curve shifts. The statement $p_{1i} \geq s > p_{0i}$ means that the instrument pushes the price from being less than s to being at least s.[7] Conditional averaging across

[7]Because p_{1i} and p_{0i} have continuous support, the statement $p_{1i} \geq s > p_{0i}$ is equivalent to the statement $p_{1i} > s > p_{0i}$.

demand curves includes only those periods where this occurs. Second, the IV estimate reflects weighted averaging along the length of a nonlinear demand curve as price varies. The weights, $\omega(s)$, are proportional to how likely each price is to be bracketed by instrument-induced price changes. The probability of bracketing is given by $\Pr(p_{1t} \geq s > p_{0t})$. Points never bracketed by p_{1t} and p_{0t} are not reflected in IV estimates. Section 3.3 develops this interpretation of Proposition 1 further using examples.

Given Assumption 1 (stationarity and ergodicity), time series data on prices can be used to estimate $\omega(p)$ as a function of p and to test whether the CDF of prices is affected by the instrument. We also note that the weighting function can be used to provide a partial check on Assumption 3 (monotonicity.) As noted above, when assumption 3 holds, $\omega(p)$ must be positive. In other words, the CDF of prices must be ordered by values of the instrument.

3.2 Sufficient conditions for monotonicity

The monotonicity condition applies to the reduced form for prices. To illustrate what this condition means for the structural equations, suppose first that the supply and demand curves are actually linear with constant coefficients and additive error terms:

$$q_t^d(p) = \alpha_0 + \alpha_1 p + \epsilon_t$$

$$q_t^s(p, z) = \beta_0 + \beta_1 p + \beta_2 z + \eta_t. \tag{8}$$

Provided, α_1 differs from β_1, the reduced form for prices is:

$$p_t = [(\beta_0-\alpha_0)/(\alpha_1-\beta_1)] + [\beta_2/(\alpha_1-\beta_1)]z_t + [(\eta_t-\epsilon_t)/(\alpha_1-\beta_1)],$$

with

$$p_{0t} = [(\beta_0-\alpha_0)/(\alpha_1-\beta_1)] + [(\eta_t-\epsilon_t)/(\alpha_1-\beta_1)],$$

$$p_{1t} = [(\beta_0-\alpha_0)/(\alpha_1-\beta_1)] + [\beta_2/(\alpha_1-\beta_1)] + [(\eta_t-\epsilon_t)/(\alpha_1-\beta_1)].$$

This linear system necessarily satisfies monotonicity because β_2, α_1, and β_1 are constant parameters. Thus, if $p_{1t} > p_{0t}$ for any t (the price is higher on day t if it has been stormy), then $p_{1t} \geq p_{0t}$ for all t, and if $p_{1t} < p_{0t}$ for any t, then $p_{1t} \leq p_{0t}$ for all t. The usual scenario for a supply and demand system is that demand slopes downwards ($\alpha_1 < 0$) and supply slopes upwards ($\beta_1 > 0$), which guarantees that p_{1t} and p_{0t} are unique. In that case, a negative supply shock ($\beta_2 < 0$) means $p_{1t} > p_{0t}$ for all t. A positive supply shock means $p_{1t} < p_{0t}$ for all t.

The following proposition provides sufficient conditions for a general time-varying, nonlinear system to satisfy monotonicity with $p_{1t} \geq p_{0t}$:

Proposition 2. $p_{1t} \geq p_{0t}$ if:

a. For all p between p_{0t} and p_{1t}, $\partial q_t^d(p)/\partial p < 0$ and $\partial q_t^s(p, 1)/\partial p > 0$, that is, demand is downward-sloping and supply is upward-sloping.

b. $q_t^s(p_{0t}, 1) \leq q_t^s(p_{0t}, 0)$, that is, z_t always indicates a negative supply shock at p_{0t}.

Proof. For some numbers p_{at} and p_{*t}, both between p_{1t} and p_{0t}, we have:

$$q_t^d(p_{1t}) = q_t^d(p_{0t}) + (\partial q_t^d(p_{at})/\partial p)(p_{1t} - p_{0t}) \tag{9a}$$

$$q_t^s(p_{1t}, 1) = q_t^s(p_{0t}, 1) + (\partial q_t^s(p_{*t}, 1)/\partial p)(p_{1t} - p_{0t}), \tag{9b}$$

by the mean value theorem. Also p_{1t} and p_{0t} are defined so that

$$q_t^d(p_{1t}) - q_t^s(p_{1t}, 1) = 0 \tag{10a}$$

$$q_t^d(p_{0t}) - q_t^s(p_{0t}, 0) = 0. \tag{10b}$$

Therefore, using (9a) to substitute for $q_t^d(p_{1t})$ in (10a), using (9b) to substitute for $q_t^s(p_{1t}, 1)$ in (10a), and subtracting (10b) from (10a), we have

$$(\partial q_t^d(p_{at})/\partial p)(p_{1t} - p_{0t}) - [q_t^s(p_{0t}, 1)-q_t^s(p_{0t}, 0)] - (\partial q_t^s(p_{*t}, 1)/\partial p)(p_{1t} - p_{0t}) = 0,$$

or

$$[(\partial q_t^d(p_{dt})/\partial p) - (\partial q_t^s(p_{st}, 1)/\partial p)](p_{1t} - p_{0t}) = q_t^s(p_{0t}, 1) - q_t^s(p_{0t}, 0).$$

This implies $p_{1t} \geq p_{0t}$ because $[q_t^s(p_{0t}, 1) - q_t^s(p_{0t}, 0)]$ and $[\partial q_t^d(p_{dt})/\partial p - \partial q_t^s(p_{st}, 1)/\partial p$ are both negative given the conditions of the proposition. □

The conditions required for Proposition 2 seem reasonable in the context of a market model. In particular, a researcher may often be prepared to assume that demand curves slope downwards and supply curves slope upwards.[8] The notion of an unambiguous supply shock also seems uncontroversial, at least in this context where supply shocks are due to bad weather. Finally, we note that Proposition 2 gives sufficient but not necessary conditions for $\alpha_{1,0}$ to be a meaningful average derivative. Violations of monotonicity need not be disastrous for IV estimates. (Monotonicity is obviously not required for identification if $\partial q_t^d(s)/\partial s$ is constant for all s and t.) In general, the bias of IV estimates involves a trade-off between heterogeneity of the response function and violations of monotonicity in the reduced form. This trade-off is explored in an evaluation example by Angrist, Imbens, and Rubin (1995.)

3.3 Examples

Two examples are used to illustrate the role played both by nonlinearity and (time-varying) heterogeneity in determining the nature of $\alpha_{1,0}$ under Proposition 1. In particular, we discuss how $\alpha_{1,0}$ is related to the population average demand function, $E[q_t^d(p)]$, and its derivative. First, suppose that potential demand is given by a non-linear but fixed functional

[8]This point motivates the approach in Manski, 1995.

relationship with the only source of time-variation coming from an additive error term (as in Roehrig, 1988, and Figure 1c):

$$q_t^d(p) = q^d(p) + \epsilon_t \tag{11}$$

In this case, the average demand function is $E[q_t^d(p)] = q^d(p)$. We have:

$$\alpha_{1,0} = \int_0^\infty [\partial q^d(s)/\partial s]\omega(s)\, ds. \tag{12}$$

Thus, when the demand function can be written as invariant across periods except for an additive error term, $\alpha_{1,0}$ is a weighted average of the slope of the average demand function, $\partial q^d(s)/\partial s$. In this case, the demand slope at *each* price involves no averaging because it is assumed to be the same each period. But the weights ranging over s reflect how likely the instruments are to shift the price from below s to at least s.

Now suppose that potential demand is linear but time-varying:

$$q_t^d(p) = \alpha_{0t} + \alpha_{1t}p + \epsilon_t$$

The average demand function for this model is $E[q_t^d(p)] = E[\alpha_{0t}] + E[\alpha_{1t}]p$. As noted by Stoker (1993), such a model could arise as the consequence of aggregating linear demand functions that differ across groups in the population (e.g., Asian and white buyers at the Fulton fish market, who appear to have different demand elasticities; see Graddy 1995.) Changes in the average demand slope over time would then be attributable to changes in the distribution of the characteristics of individuals coming to market.

It is easiest to simplify $\alpha_{1,0}$ in this case using Lemma 1:

$$E[q_t|\ z_t=1] - E[q_t|\ z_t=0] = E\{ \int_{p_{0t}}^{p_{1t}} [\partial q_t^d(s)/\partial s]\, ds \} = E[\alpha_{1t}(p_{1t} - p_{0t})]$$

Therefore,

$$\alpha_{1,0} = E[\alpha_{1t}(p_{1t} - p_{0t})]/E[p_{1t} - p_{0t}].$$

This is generally not equal to the slope of the average demand curve, $E[\alpha_{1t}]$. Note that in this case, there is averaging across periods even though the within-period demand slope is assumed to be constant.

In this example, periods when p_{1t}-p_{0t} is largest are likely to be over-represented in $\alpha_{1,0}$. If such periods are not characteristic of the demand relationship in the future, perhaps because groups most affected by the instrument are not particularly important in the aggregate demand relationship, then $\alpha_{1,0}$ may be of little value for predicting the impact of exogenous price changes in the future. Of course, if α_{1t} is independent of $(p_{1t} - p_{0t})$, then $\alpha_{1,0} = E[\alpha_{1t}]$. This is unlikely, however, unless α_{1t} is actually constant. Note that for any linear demand equation:

$$\alpha_{1,0} = E[\alpha_{1t}(p_{1t} - p_{0t})]/E[p_{1t} - p_{0t}] = E[\alpha_{1t}] + Cov[\alpha_{1t}, p_{1t} - p_{0t}]/E[p_{1t} - p_{0t}]. \qquad (13)$$

If the supply curve is also linear, as in equation (8), then $(p_{1t} - p_{0t}) = \beta_2/(\alpha_{1t} - \beta_1)$ and the previous equation simplifies to:

$$\alpha_{1,0} = E[\alpha_{1t}] + \beta_2 Cov[\alpha_{1t}, (\alpha_{1t} - \beta_1)^{-1}]/E[p_{1t} - p_{0t}] > E[\alpha_{1t}].$$

Note that $\beta_2 < 0$, $Cov[\alpha_{1t}, (\alpha_{1t} - \beta_1)^{-1}] < 0$ and $\alpha_{1t} < 0$. Therefore, $\alpha_{1,0}$ is smaller in magnitude than the average elasticity, $E[\alpha_{1t}]$. The intuition for this is that when $|\alpha_{1t}|$ is high (demand is elastic), the price gap p_{1t}-p_{0t}, which weights each estimate, is small. $\alpha_{1,0}$ therefore weights high elasticity periods less than $E[\alpha_{1t}]$.[9]

The two examples in this subsection illustrate two features which link $\alpha_{1,0}$ to the source

[9] A similar link between an underlying behavioral model and the interpretation of IV estimates is made by Card (1994) in a survey of IV estimates of the returns to schooling.

of variation captured by the instrument. The first example shows that when the underlying response function is nonlinear, more weight is given to points on the response function where $Pr[p_{1t} \geq s > p_{0t}]$ is large. These are points that are most likely to be bracketed by an instrument-induced shift in prices. The second example shows that even if the underlying response function is linear, the weight given to the slope coefficient for any period is proportional to the difference between p_{1t} and p_{0t} in that period. Thus, $\alpha_{1,0}$ is not necessarily useful for characterizing the impact of an exogenous price change that is rarely bracketed by $[p_{0t}, p_{1t}]$ or expected to occur in an environment where $p_{1t}-p_{0t}$ is small.

Whether or not the unique features of $\alpha_{1,0}$ are restrictive must be considered on a case by case basis. For example, if a researcher is prepared to assume that the demand relationship is nonlinear but unchanging except for additive shocks, then $\alpha_{1,0}$ should prove useful for forecasting the impact of exogenous price changes in a range that is commonly between p_{1t} and p_{0t}. A useful estimate of this range is likely to be given by $\{E[p_{0t}], E[p_{1t}]\}$. We also note that the quality of an instrument i.e., the magnitude of the first-stage relationship) affects this range.

4. Continuous instruments and discrete covariates

This section discusses estimates computed using continuous instruments of the following type:

Assumption R2. z has continuous support on $[0, \infty)$ and $E[p_t| z]$ is a continuously differentiable function of z.

Assuming non-negative z is convenient for some of the derivations that follow and involves no loss of generality because we can always work with a transformation that has this property (e.g., exp[z].) We continue to assume that the only element of x_t is a constant.

The reduced form associated with a continuous instrument is derived from the potential excess demand function, $e_t(p, z)$, defined as:

$$q_t^d(p, z) - q_t^s(p, z) \equiv e_t(p, z). \tag{13}$$

We also impose some technical restrictions:

Assumption R3. For all values in the support of (p, z) and for all t:

a. $q_t^d(p, z)$ and $q_t^s(p, z)$ are continuously differentiable in z.

b. $\partial e_t/\partial p \neq 0$.

Note that continuous differentiability of $E[q_t | z]$ is implied by assumption R2 in combination with R3. Given R3, the potential excess demand function determines an equilibrium price function in period t, $h_t(z)$, satisfying:

$$e_t[h_t(z), z] = 0.$$

The function $h_t(z)$ is the implicit reduced form. For observed prices and quantities, we have:

$$p_t = h_t(z_t)$$

$$q_t = q_t^d(h_t[z_t]) \equiv g_t(z_t)$$

Therefore,

$$E[p_t | z_t=z] \equiv E[p_t | z] = E[h_t(z)]$$

$$E[q_t | z_t=z] \equiv E[q_t | z] = E[q_t^d(h_t[z])] \tag{14}$$

We emphasize that, as with the structural equations, $h_t(z)$ and $g_t(z)$ are random variables giving potential prices and quantities for every possible value of z in period t.

Finally, we define monotonicity of the reduced form for this case:

Assumption 4 (monotonicity for continuous instruments.) At each possible value of the instrument, z, either $h_t'(z) \geq 0$ for all t or $h_t'(z) \leq 0$ for all t.

Assumption 4 restricts the sign of the derivative of the reduced form at values of z other than the observed z_t. This extends Assumption 3 to the case of continuous instruments and has the same testable implications. Like Assumption 3, it cannot be tested directly because $h_t'(z)$ is observed only for $z=z_t$. Note that monotonicity requires only that at each z, *either* $h_t'(z) \geq 0$ for all t or $h_t'(z) \leq 0$ for all t. But this implies that the instrument can be re-ordered and redefined so that $h_t'(z) \geq 0$ for all t and z. For the purposes of the next section, we assume this has been done.

4.1 Use of $E[p_t | z_t]$ as an instrument

This estimator is efficient for a homoscedastic regression model in the sense that it attains the efficiency bound for conditional moments estimation of the constant-coefficients homoscedastic model. In practice, $E[p_t | z_t]$ is unknown but can be estimated parametrically or nonparametrically (see, e.g., Newey 1990). Estimation of the reduced form does have implications for the calculation of standard errors. Formulas for asymptotic standard errors of IV estimates under the assumptions in this paper are developed in the appendix.

The estimator of interest in this section is:

$$\alpha_B = \frac{E\{ q_t[E(p_t | z_t) - \mu_p] \}}{E\{ p_t[E(p_t | z_t) - \mu_p] \}},$$

where μ_p is the mean of p_t. At this point, we define the analog of a Wald-type parameter for a continuous instrument, defined at each value of the instrument:

$$\alpha(z) = \lim_{\nu \to 0} \frac{E[q_t | z] - E[q_t | z-\nu]}{E[p_t | z] - E[p_t | z-\nu]} .$$

This is the population Wald estimate associated with a small change in z. It is related to the slopes of the demand curve and reduced form by the following lemma:

Lemma 2. $\alpha(z) = (\partial E[q_t | z]/\partial z)/(\partial E[p_t | z]/\partial z) = E\{ [\partial q_t^d(h_t[z])/\partial p)] \cdot [h_t'(z)] \}/E[h_t'(z)]$.

Proof. The first equality is immediate from the definition of a derivative, the second equality is a consequence of equation (14). □

Thus, $\alpha(z)$ is a weighted average of the slope of the demand curve at price $h_t(z)$. The weights in this case range over the distribution of $h_t(z)$ for fixed z, and are given by the derivative of the reduced form relationship (which varies with t). Note that $\alpha(z)$ is not defined for values of z at which $E[h_t'(z)]=0$. Also, periods in which $h_t'(z)=0$ do not contribute to $\alpha(z)$.

Proposition 3 characterizes α_E:

Proposition 3. $$\alpha_E = \int_0^\infty \alpha(s)\, \omega(s)\, ds,$$

where

$$\omega(s) = Var\{E[p_t | s]\}^{-1} \times$$

$$\{(E[p_t | z_t \geq s] - E[p_t | z_t < s]) \cdot Pr[z_t \geq s](1-Pr[z_t \geq s]) \cdot E[h_t'(s)]\},$$

and where s is an integrating variable for z, with $\alpha(s) = (\partial E[q_t | s]/\partial s)/(\partial E[p_t | s]/\partial s)$, as defined in Lemma 2. The weight function, $\omega(s)$, is non-negative and integrates to 1.

Proof. $E[q_t| z_t] = E[q_t| z=0] + \int_0^{z_t} (\partial E[q_t| s]/\partial s)\, ds.$

Therefore, $E\{ q_t[E(p_t| z_t) - \mu_p] \} = E\{ E(q_t| z_t)[E(p_t| z_t) - \mu_p] \} =$

$$E\left\{ \left(\int_0^{z_t} \partial E[q_t| s]/\partial s\, ds\right) (E[p_t| z_t] - \mu_p) \right\}$$

$$= \int_0^{\infty} \left(\int_0^{z_t} \partial E[q_t| s]/\partial s\, ds\right) (E[p_t| z_t] - \mu_p)\, r(z_t)dz_t$$

where $r(z_t)$ is the density of z_t. Changing the order of integration gives

$$= \int_0^{\infty} \int_s^{\infty} (E[p_t| z_t] - \mu_p)\, r(z_t)dz_t\, (\partial E[q_t| s]/\partial s)\, ds$$

$$= \int_0^{\infty} (E[p_t| z_t \geq s] - \mu_p)\Pr(z_t \geq s)(\partial E[q_t| s]/\partial s)\, ds.$$

Using the fact that $\mu_p = E[p_t| z_t \geq s]\Pr(z_t \geq s) + E[p_t| z_t < s][1-\Pr(z_t \geq s)]$, we have $(E[p_t| z_t \geq s] - \mu_p)\Pr(z_t \geq s) = (E[p_t| z_t \geq s] - E[p_t| z_t < s])\Pr[z_t \geq s](1-\Pr[z_t \geq s])$. Finally, using the previous lemma to substitute for $\partial E[q_t| s]/\partial s$ completes the derivation of the numerator of $\omega(s)$. An argument similar to this can also be used to show that the numerator integrates to $\mathrm{Var}\{E[p_t| z]\}$, thereby proving that the weighting function integrates to 1. The numerator is non-negative because $\Pr(z_t \geq s) > 0$ and because monotonicity implies $h_t'(s) \geq 0$ and $(E[p_t| z_t \geq s] - E[p_t| z_t < s]) \geq 0$. □

Proposition 3 shows that IV estimation using $E[p_t| z_t]$ as an instrument produces a weighted average of average causal derivatives calculated at each value of the instrument.

Values of the instrument getting more weight are those close to the median (where $\Pr[z_i \geq s](1-\Pr[z_i \geq s])$ is largest) and those where the instrument is most strongly related to the price. In particular, values of z_i where the reduced form has a slope of zero do not get any weight.

Finally, we note that sufficient conditions for monotonicity to hold in the continuous-instrument case can be obtained directly from the implicit function theorem. These conditions are similar to those in Proposition 2, with the modification that restrictions on the supply shift are cast in terms of the derivative of the supply curve with respect to z.

4.2 Incorporating covariates

Conditional on discrete covariates, the problem of interpreting IV estimates is as outlined above. For example, the primary covariate of interest in our application is day of the week. Although weather conditions do not vary systematically with day of the week, there is substantial residual variation in quantities and prices, perhaps due to demand fluctuations, that can be reduced by conditioning on day of the week. Controlling for day to day variability could therefore lead to more efficient estimates of the parameter of primary interest. Non-parametric conditioning on discrete covariates is one approach discussed by Stoker (1986, p. 1470) in his work on average derivative estimation. Examples of IV applications where all covariates are discrete include Angrist (1990), Angrist and Krueger (1991, 1992), and Imbens and van der Klaauw (1995).

When covariates take on many values, non-parametric conditioning can lead to many imprecise estimates. It is therefore of interest to consider the probability limit of the instrumental variables estimator in models that allow for a changing intercept but fix the price coefficient across covariates. The following is adapted from Angrist and Imbens (1995):

Proposition 4. Let $D[x_t]$ be the t^{th} row from a design matrix constructed from indicator variables for each value of a discrete covariate, x_t. Consider the 2SLS estimate computed using $E[p_t| x_t, z_t]$ as an instrument for p_t in a regression of q_t on $D[x_t]$ and p_t. The estimator for the coefficient on p_t is:

$$\alpha_{E|x} = \frac{E\{q_t(E[p_t| x_t, z_t] - E[p_t| x_t])\}}{E\{p_t(E[p_t| x_t, z_t] - E[p_t| x_t])\}} \qquad (15)$$

$$= \frac{E[\alpha(x_t)\theta(x_t)]}{E[\theta(x_t)]}$$

where $\theta(x_t) = E\{E[p_t| x_t, z_t](E[p_t| x_t, z_t] - E[p_t| x_t])| x_t\}$ and

$$\alpha(x_t) = \frac{E\{q_t(E[p_t| x_t, z_t] - E[p_t| x_t])| x_t\}}{E\{p_t(E[p_t| x_t, z_t] - E[p_t| x_t])| x_t\}}.$$

Proof. Equation (15) is immediate from the definition of 2SLS and the fact that the reduced form is a saturated model for discrete regressors. The weighting formula can be established by iterating expectations. □

Note that $\alpha(x_t)$ is the 2SLS estimate constructed using $E[p_t| x_t, z_t]$ as an instrument in a population where x_t is fixed. Proposition 4 therefore says that 2SLS estimates of the coefficient on a single endogenous regressor in a model with dummy variable covariates is a weighted average of the 2SLS estimates conditional on the covariates. The weights consist of the variance of $E[p_t| x_t, z_t]$ conditional on the covariates. In some cases, $E[p_t| x_t, z_t]$ may be wellapproximated by a linear regression on $D[x_i]$ and z_1. In such cases, the estimator described in the proposition corresponds to 2SLS without interaction terms in the instrument list.

4.3 Relationship to average-derivative estimators

Stoker (1986) and Powell, Stock and Stoker (1989) also explore strategies for the nonparametric estimation of average derivatives. Stoker (1986) shows that for a conditional expectation function $E[y|x] = g(x)$, where x has density f(x), using $-f'(x)/f(x)$ as an instrument generates a consistent estimate of the average derivative of g(x). Similarly, Powell, Stock, and Stoker (1989) develop estimators based on the fact that,

$$\mathrm{Cov}[f'(x),y]/\mathrm{Cov}[f'(x),x] \;=\; E[f(x)(\partial g/\partial x)]/E[f(x)].$$

In other words, the density-weighted average derivative of an unknown regression function can be obtained by using the derivative of the density of regressors as an instrument for a regression of y on x.

Powell, Stock and Stoker (1989) focus on kernel estimation of the unknown density derivative, $f'(x)$, and on obtaining standard errors for the resulting instrumental variables estimate. The results in these papers are principally motivated by a desire to estimate index coefficients since, if $g(x) = G(x'\kappa)$ for some function $G(\cdot)$ and coefficient vector κ, any weighted average derivative is proportional to κ. Neither paper is concerned with models involving endogenous regressors. An implication of our results is that the average-derivative property underlying the Stoker and Powell, Stock, and Stoker papers is a general feature of IV estimates. In this case, however, our interest in average derivatives is not tied to an underlying index model. Rather, average derivatives are viewed here as a causal parameter for an unknown heterogeneous response function involving endogenous regressors.

5. Application: The demand for whiting at the Fulton fish market

We illustrate the average-derivative interpretation of IV estimates by estimating the average elasticity of wholesale demand for fresh whiting at the Fulton fish market in New York City. This application is useful for our purposes because strong and credible instruments are available in the form of weather conditions at sea. The quantity of fish brought to market is determined by many factors, but the weather is a primary force in influencing how much fish will arrive since wind and waves make it difficult to catch fish.

Fish is sold by about 35 different dealers at the Fulton market, although only six of the dealers regularly sell whiting. There are no posted prices, and each dealer is free to charge a different price to each customer. Dealers can leave the Fulton market and new dealers can rent stalls, although in practice this happens rarely and did not happen over the sample period. The buyers at the Fulton market generally own retail fish shops or restaurants.

Whiting is a good choice for a study of the wholesale fish market because more transactions take place in whiting than almost any other fish. Whiting also vary less in size and quality than other fish. Finally, there is probably very little substitution between whiting and other fish. Whiting is a very cheap fish in large supply that is oily and distinctive tasting. Other fish would rarely be sold at a low enough price and in sufficient quantities to be attractive to retailers or restaurants as a substitute for whiting.

The data used in Graddy (1995) were obtained from a single dealer who supplied his inventory sheets for the period December 2, 1991 through May 8, 1992. Total price and quantity for each transaction are recorded on the inventory sheets. These data are supplemented by data that were collected by direct observation from the same dealer during the period April 13 through May 8, 1992. For this study, the prices and quantities are aggregated by day, for the 111 days the market was open between December 2 and May 8. The price variable used below is the average daily transaction price for the dealer observed. The quantity variable is the total quantity sold by this dealer on each day.

5.1 The nature of the wholesale demand for fish

Every day the demand for fish at the Fulton Fish market is determined partly by which customers have decided to visit the market that day, as well as by how much they buy. A number of customers visit the market every week on Mondays and Thursdays, and other customers may visit the market every day of the week. Quantities purchased by individual wholesale customers with potentially different demand functions are summed up to produce the daily aggregate demand for fish. Aggregation is therefore one source of time-series shifts in potential demand. For example, Graddy (1995) presents evidence which suggests that Asian buyers have a more elastic demand for fish than whites and that the ethnic mix of buyers changes from day to day.

We begin our analysis with a demand function for the quantity demanded by customer c on day t:

$$q^d_{tc}(\mathit{ln}p_{tc}, x_t).$$

The subscript c ranges over the list of customers who ever visited the Fulton Fish market during our sample period (c = 1, . . ., C). We define potential demands in terms of lnp so that price changes are in percentage terms. As before, x_t is a vector of covariates shifting demand.

To distinguish between the average daily price and the price paid by a particular customer at a particular time, we write

$$lnp_{tc} = lnp_t + ln(p_{tc}/p_t) \equiv lnp_t + \nu_{tc},$$

where p_t is the average price on day t. Total quantity demanded is the sum of customer demand (including zeros for customers who make no purchases on a given day, for whom p_{tc} is the price that would have been quoted if they had gone to market at the usual time):

$$Q_t(lnp_t, \nu_{t1}, \ldots, \nu_{tC}; x_t) \equiv \Sigma_c\, q^d_{tc}(lnp_t + \nu_{tc}; x_t).$$

The potential demand at average log price lnp is defined as

$$q_t(lnp, x_t) \equiv Q_t(lnp, \nu_{t1}, \ldots, \nu_{tC}; x_t).$$

Thus, $q_t(lnp, x_t)$ is the sum of customer demands at some hypothetical average log price, given the same intra-day, intra-customer dispersion in prices.

Assumption 2 requires that $q_t(lnp, x_t)$ be independent of the instruments. We must therefore assume that given p_t, who comes to market and what time they make their purchase is independent of weather conditions at sea. To clarify the nature of this assumption, suppose that on days in which the weather forecast at sea is particularly bad, customers arrive at the market earlier than usual. Since arrival times are associated with differences in price quotes (see Graddy, 1995), such behavior could lead to a violation of the assumptions underlying the use of the average daily price as a "sufficient statistic" for quantity demanded.

The object of estimation (the estimand) is the IV estimate of α in the following equation:

$$lnq_t = \gamma_o + X_t'\gamma_1 + \alpha lnp_t + \xi_t \qquad (16)$$

where q_t is the quantity sold on a particular date, γ_o is a constant, X_t is a vector of dummy variables for different days of the week, γ_1 a vector of coefficients, and p_t is the average price at date t. The interpretation of the estimated α is given by Proposition 1 or Proposition 3. Thus, ξ_t is *defined* as $(lnq_t - \gamma_0 - X_t'\gamma_1 - \alpha lnp_t)$, where γ_0, γ_1 and α are the probability limits of IV estimates. ξ_t is not a structural error term or the deviation from a hypothetical conditional expectation. The double-log format is chosen for (16) so that the IV estimates will be immediately interpretable as an average elasticity. We emphasize that equation (16) is a computational device, and not an assumption about the shape of demand functions.

Including dummy-variable covariates gives the IV slope estimate the interpretation of a variance-weighted average over days. In practice, the empirical results show that the estimates are not affected by inclusion of these dummies, suggesting fairly homogeneous demand from day to day. A second set of important covariates in demand analysis is the price of substitute products, in this case, the price of other types of fish. Such prices are not included as covariates because of the apparent lack of substitution between whiting and other types of fish.

As noted above, equation (16) is the equation we would like to estimate. In practice, we only have data on prices and quantities for a single dealer, out of the 6 dealers that carried whiting at the Fulton Fish market at the time Graddy's (1995) study was carried out. Let q_{dt} denote quantity sold by the dealer providing information. By definition,

$$lnq_{dt} = lnq_t + lns_{dt} \tag{17}$$

where s_{dt} is the reporting dealer's market share. Provided weather-related supply shocks affect all dealers equally, IV estimates using lnq_{dt} as a dependent variable will be the same as those using lnq_t as a dependent variable. Such an assumption seems innocuous but is clearly essential.

5.2 Reduced form and IV estimates

Table 1 reports reduced form estimates of the relationship between prices and quantities and weather conditions. The dummy instrument, Stormy, indicates wave height greater than 4.5 feet and wind speed greater than 18 knots. A second dummy instrument, Mixed, indicates wave height greater than 3.8 feet and wind speed greater than 13 knots if Stormy equals zero. Wind speed and wave height are moving averages of the last three days' wind speed and wave height before the trading day, as measured off the coast of Long Island and reported in the New York Times boating forecast.

The reduced form results show that Stormy is a statistically significant determinant of both the price and quantity of fish sold at the Fulton market. Stormy weather decreases the quantity and increases the price. The covariates are not significant in the price equation, but they do show that the quantity purchased on Tuesday and Wednesday is significantly less than that purchased on other days of the week. The coefficients on Mixed and Stormy are increasing in absolute value, but the coefficients on Mixed in the quantity equations are not significant.

Columns (4-6) of Table 2 report IV estimates using Stormy as an instrument, and columns (7-9) report 2SLS estimates using both Stormy and Mixed as instruments. For comparison, columns (1-3) report the corresponding OLS estimates. IV and 2SLS estimates of the price coefficient are almost twice as large as the OLS estimates (with much larger standard errors as well). The price coefficients are all statistically significant, and change very little when exogenous covariates are included in the regressions. This reflects the fact that the instruments are essentially orthogonal to the covariates. The estimated standard errors are slightly smaller when two binary instruments are used instead of just one. The fact that the 2SLS estimates are typically a bit larger than one is noteworthy because profit-maximization with an element of collusive behavior on the part of dealers should generate transactions in a range where demand is more than unit elastic (see, e.g., Bresnahan, 1989).

Table 2 also reports estimates of equations that include controls for shore weather conditions in case weather conditions at sea are not a valid instrument by virtue of correlation with weather conditions on shore. The variables characterizing shore weather are Cold, indicating temperature less than the median average daily temperature (41.5 degrees) and Rainy, indicating any precipitation on day t. Including Rainy and Cold in the estimation has remarkably little effect on the price coefficients. These variables also appear unrelated to market volume.

5.3 The weighting function for a binary instrument

The weighting function for a binary instrument is the difference in the CDF for prices conditional on values of the instrument. We use the sample distribution function, $S_T(p)$, to estimate the weighting function. $S_T(p)$ is defined as the proportion of observations that are less than or equal to p. The sample CDFs of price for stormy and clear weather conditions are plotted in Figure 2. As required by the monotonicity assumption, the CDFs do not cross. The difference in the two sample CDFs, normalized to sum to one, is plotted in Figure 3. This is the weighting function underlying IV estimates using Stormy as an instrument. Figure 3 also shows the histogram of prices. A one-sided Kolmogorov-Smirnov test (see, e.g., Gibbons 1985) rejects equality of the price distributions at the 5% level.

The CDF difference exhibits a sharp peak at the point at which price per pound is about \$1.00. The mean price per pound over all days is \$0.88 and the median is \$0.81, suggesting that days when prices were higher are contributing disproportionately to IV estimates based on Stormy. Although Stormy provides some information on demand response over prices ranging from \$0.50 to \$1.50, Figure 3 shows that the primary effect of Stormy on prices occurs in the second and third price quartiles. These are days in which prices ranged from \$0.62 to \$1.19. If the cutoff point for Stormy were at lower wind speeds and wave heights, days with lower prices would contribute more to the IV estimates. Finally, note that comparing the IV weighting function to the price histogram, we get an idea how a density-weighted average of the Powell, Stock, and Stoker (1989) type would capture different parts of the demand function. Of course, computation of a density-weighted average requires exogenous prices.

Figure 4 graphs the sample CDF of prices in stormy, mixed, and fair weather conditions. The CDFs are clearly ordered by weather severity, suggesting that monotonicity is satisfied for IV estimates based on any pairwise comparison of prices and quantities in these 3 weather conditions. 2SLS estimates using both Stormy and Mixed dummies as instruments can be interpreted as a linear combination of such pairwise Wald estimates (Angrist and Imbens 1995.)

5.4 Instrumental Variables estimates using $E[p_t | z_t]$

Figure 5 plots log price against the log of a 3-day moving average of wind speed.[10] Also shown are the fitted values from a quadratic regression of log price on log wind speed. The log-quadratic appears to capture essential features of the relationship between weather severity and log prices. We therefore take the log-quadratic to be an estimate of $E[\mathit{ln}p_t | \mathit{ln}z_t]$, where z_t is

[10]Wind speed is reported in 5-knot intervals. In the 110-day sample used here, the wind speed moving average takes on 21 different values.

wind speed.

Columns (10-12) of Table 2 report 2SLS estimates of the demand equation using linear and quadratic terms in lnz_t as instruments. This is equivalent to using an OLS estimate of $E[lnp_t| \; lnz_t]$ as a single instrument. The weighting function underlying the continuous IV estimates is graphed in Figure 6 as a function of log wind speed. The denominator of the weighting function is the variance of the fitted values. Components of the numerator are also computed from the quadratic fit.[11] Recall that for every value of a continuous instrument, there is an "instantaneous Wald estimate" giving the average causal response induced by a small change in the instrument at that point (Lemma 2.) The weighting function in Figure 6 shows which and how much each of these Wald estimates are contributing to IV estimates based on $E[lnp_t| \; z_t]$. The histogram of log wind speed is also shown for comparison.

Note that the weighting function in Figure 6 is conceptually different from the weighting function in Figure 3, which refers to a binary comparison. In fact, for *each* of the pointwise Wald estimates underlying estimates based on $E[lnp_t| \; z_t]$, it is theoretically possible to evaluate a weighting function like that in Figure 3, which shows how a binary comparison traces out part of the demand function. This weighting function would involve the derivative with respect to z_t of the CDF of prices conditional on z_t. In contrast, the weighting function in Figure 6 provides information on the relative importance of different values of a continuous instrument. The fact that IV and 2SLS estimates in Table 2 are similar suggests that the response of quantity demanded is close to unit elastic for all weather-induced price variation in our sample.

---

[11]In particular, for each value of the instrument, s, $E[p_t| \; z_t \geq s]$ is estimated using predicted prices from the quadratic reduced form, and $E[h_t'(s)]$ is the derivative of the quadratic fit evaluated at s.

6. Conclusions

The interpretation of the SEM in this paper appears to be at odds with the notion that structural modelling in economics should be directed towards identifying behavioral parameters that are invariant across populations. For example, Lucas (1976) defines structural parameters in econometric models as those that are invariant under hypothetical interventions in the process generating outcomes. In contrast to this emphasis on invariance, Goldberger (1972) has defined structural equations models as those "in which each equation represents a causal link, rather than a mere empirical association" (p. 979).

This paper is motivated by the view that IV estimates of demand equations are interesting because they tell us something about a causal or behavioral link, and not necessarily because they are invariant. Causal relationships need not be linear, parallel, or invariant across periods or markets. Moreover, causal estimates, like any statistical estimates, are necessarily tied to the source of variation generating the estimates. In particular, our discussion of simultaneous equations models highlights the fact that instrumental variables estimates, while capturing important aspects of a behavioral relationship, are not necessarily structural in the Lucas sense. On the other hand, combined with additional evidence that circumstances have not changed in essential ways, IV estimates may be useful for forecasting the impact of future interventions. The question of how to use IV estimates most effectively for policy-making is a natural avenue for future research.

References

Amemiya, Takeshi, 1974, "The Nonlinear Two-Stage Least Squares Estimator," *Journal of Econometrics* 2, 105-110.

Angrist, J., 1990, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *American Economic Review*, 80, 313-335.

________ and A. Krueger, 1991, "Does Compulsory School Attendance Affect Schooling and Earnings?", *Quarterly Journal of Economics*, 106, 979-1014.

________ and ________, 1992, "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association* 87, 328-336.

Angrist, J.D., and G.W. Imbens, 1995, "Two-Stage Least Squares Estimation of Average Causal Response in Models with Variable Treatment Intensity," forthcoming, *Journal of the American Statistical Association* 90.

________, ________ and D.B. Rubin, 1995, "Identification of Causal Effects Using Instrumental Variables", forthcoming, *Journal of the American Statistical Association* 90.

Bresnahan, T.F., 1989, "Empirical Studies of Industries with Market Power, in R. Schmalensee and R.D. Willig, eds., *The Handbook of Industrial Organization*," New York: North-Holland, 1989.

Card, D.E., 1994, "Earnings, Schooling, and Ability Revisited," Industrial Relations Section, Princeton University, Working Paper No. 331, May.

Gibbons, J.D., 1985, *Nonparametric Methods for Quantitative Analysis*, 2nd Edition, Columbus: American Sciences Press.

Goldberger, A.S., 1972, "Structural Equations Models in the Social Sciences," *Econometrica* 40, 979-1001.

Graddy, Kathryn, 1995, "Testing for Imperfect Competition at the Fulton Fish Market," forthcoming, *Rand Journal of Economics*.

Haavelmo, Trygve, 1944, *The Probability Approach in Econometrics*, Supplement to *Econometrica* 12, July 1944, Chicago: The Econometric Society.

Hausman, J.A., and W.K. Newey, 1995, "Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss," *Econometrica*, forthcoming.

Heckman, J.J., 1992, "Haavelmo and the Birth of Modern Econometrics: A Review of *The History of Econometric Ideas* by Mary Morgan," *Journal of Economic Literature* 30, 876-886.

Imbens, G.W., and J.D. Angrist, 1994, "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62, 467-476.

________, and D.B. Rubin, 1994a, "The Fragility of Instrumental Variables Estimators," Harvard Institute for Economic Research Working paper No. 1675.

________, and D.B. Rubin, 1994b, "Causal Inference with Instrumental Variables," Harvard Institute for Economic Research Working paper No. 1676.

________, and W. van der Klaauw, 1995, "Estimating the Cost of Conscription in the Netherlands," *Journal of Business and Economic Statistics* 13, April.

Kelejian, H., 1971, "Two-Stage Least Squares and Econometrics Systems Linear in Parameters and Nonlinear in Endogenous Variables," *Journal of the American Statistical Association* 66, 373-374.

Lucas, R.E., 1976, "Econometric Policy Evaluation: A Critique," in *The Phillips Curve and the Labor Market*, edited by K. Brunner and A. Meltzer, Volume 1 of Carnegie-Rochester Conferences on Public Policy.

Manski, Charles F., 1995, "Simultaneity With Downward Sloping Demand," Chapter 6 in *Identification Problems in the Social Sciences*, Cambridge: Harvard University Press.

Newey, W.K., 1990, "Efficient Instrumental Variables Estimation of Nonlinear Models," *Econometrica* 58, 809-838.

________, and J.L. Powell, 1989, "Instrumental Variables Estimation for Nonparametric Models," Princeton University Department of Economics, mimeo, October.

________, ________, and F. Vella, 1994, "Nonparametric Estimation of Triangular Simultaneous Equations Models," MIT Department of Economics, October.

________, and K. West, 1987, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, 703-708.

Neyman, J., 1923, "On the Application of Probability theory to Agricultural Experiments, Essay on Principles, Section 9," Translated in *Statistical Science* 5, 465-480.

Powell, J.L., J.H. Stock, and T.M. Stoker, 1989, "Semiparametric Estimation of Index Coefficients, *Econometrica* 57, 1403-1430.

Roehrig, Charles S., 1988, "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica* 56, 433-447.

Rubin, D.B, 1978, "Bayesian Inference for Causal Effects: The Role of Randomization," *Annals of Statistics*, 6, 34-58.

__________ 1990, "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies," *Statistical Science* 5, 472-480.

Stoker, T.M., 1986, "Consistent Estimation of Scaled Coefficients," *Econometrica* 54, 1461-1481.

__________, "Empirical Approaches to the Problem of Aggregation Over Individuals," *Journal of Economic Literature* 31 (December 1993), 1827-1874.

White, H., 1984, *Asymptotic Theory for Econometricians*, Cambridge: Cambridge University Press.

Appendix

This appendix describes how standard errors for the IV estimates discussed in Section 5 were calculated.[12] To simplify notation, the analysis refers to a model without covariates. In this case, the estimating equation, (16), simplifies to

(A1) $$ln q_t = \gamma_o + \alpha ln p_t + \xi_t$$

and the first-stage equation is

(A2) $$ln p_t = Z_t'\pi + u_t,$$

where π is a $k\times 1$ vector of coefficients and Z_t is a conformable vector of instruments that includes a constant. The reduced form error term satisfies $E[Z_t u_t]=0$.

In the framework of this paper, the moment conditions that identify γ_0 and α are:

(A3) $$E[\xi_t]=0$$

$$E[\pi' Z_t \xi_t]=0.$$

A key difference between this approach and conventional IV estimation is that we do not assume $E[Z_t\xi_t]=0$ because the model outlined here does not imply that each instrument leads to the same structural coefficients in the population. In other words, γ_0 and α are *defined* as the estimators that satisfy a just-identified IV problem using the 2SLS instrument, $\pi' Z_t$. The individual elements of Z_t would not necessarily generate the same coefficient estimates if used separately.

Stacking the moment conditions for the reduced form and structural equation, generates the following $(k+2)\times 1$ vector of sample moments:

---

[12] Imbens and Rubin (1995) develop a Bayesian approach to IV inference that can lead to more accurate inference than conventional asymptotic approximations.

$$(1/T)\Sigma_t \psi_t(Z_t; \gamma_0, \alpha, \pi)=0$$

where

$$\psi_t(Z_t; \gamma_0, \alpha, \pi) = \begin{bmatrix} (lnq_t - \gamma_o - \alpha lnp_t) \\ \pi' Z_t(lnq_t - \gamma_o - \alpha lnp_t) \\ Z_t(lnp_t - Z_t'\pi) \end{bmatrix}.$$

There are k+2 parameters to be estimated so that this is a just-identified GMM problem, regardless of the number of instruments. The fact that the α which satisfies the structural-equation moment condition changes with the vector of reduced form coefficients, π, means that estimation of π will generally affect the limiting distribution of α.

Using the heteroskedasticity and autocorrelation consistent covariance matrix estimator suggested by Newey and West (1987) for GMM problems of this type, we can estimate the asymptotic covariance matrix of $\theta \equiv (\gamma_0, \alpha, \pi)$ as

$$\hat{V} = (1/T)[A^{-1}BA^{-1}]$$

where

$$A = (1/T)\Sigma_t \, \partial\psi_t(\theta)/\partial\theta$$

$$B = \Omega_0 + \sum_{j=1}^{L} (1-j/(L+1))(\Omega_j + \Omega_j')$$

and

$$\Omega_j = (1/T) \sum_{t=j+1}^{T} \psi_t(\theta)\psi_{t-j}(\theta)'.$$

In practice, a lag length of L=5 was used to calculate the reported standard errors. The tables show conventional OLS and IV standard errors in parentheses and standard errors calculated as described here in square brackets. Table 2 also shows the conventional chi-square goodness-of-fit statistics for the IV moment conditions.

Table 1. Reduced Form Estimates for Quantity and Price.

	Quantity (1)	Price (2)	Quantity (3)	Price (4)	Quantity (5)	Price (6)	Quantity (7)	Price (8)
Stormy	-0.363 (0.152) [0.158]	0.335 (0.074) [0.081]	-0.388 (0.144) [0.161]	0.346 (0.075) [0.079)	-0.449 (0.168) [0.176]	0.437 (0.078) [0.097]	-0.433 (0.159) [0.181]	0.446 (0.079) [0.095]
Mixed					-0.201 (0.165) [0.178]	0.236 (0.077) [0.101]	-0.106 (0.157) [0.168]	0.237 (0.079) [0.100]
Monday			0.101 (0.207) [0.184]	-0.113 (0.107) [0.082]			0.099 (0.207) [0.188]	-0.108 (0.103) [0.087]
Tuesday			-0.485 (0.201) [0.167]	-0.041 (0.105) [0.088]			-0.474 (0.202) [0.165]	-0.066 (0.101) [0.085]
Wednesday			-0.553 (0.206) [0.164]	-0.012 (0.107) [0.091]			-0.536 (0.208) (0.169]	-0.049 (0.104) [0.084]
Thursday			0.054 (0.201) [0.173]	0.050 (0.104) [0.064]			0.058 (0.202) [0.169]	0.039 (0.101) [0.066]
no. of obs.	111	111	111	111	111	111	111	111

Notes: Standard errors in parentheses were calculated using the usual asymptotic approximation for OLS and 2SLS or IV. Standard errors in square brackets were calculated using the formulas in the appendix.

Table 2. OLS, IV and 2SLS estimates of the demand equation.

	OLS			IV (Stormy)			2SLS (Stormy, Mixed)			IV (Predicted price)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
price	-0.541	-0.563	-0.545	-1.082	-1.119	-1.223	-1.014	-0.930	-0.947	-1.230	-1.077	-1.177
	(0.179)	(0.168)	(0.175)	(0.466)	(0.429)	(0.532)	(0.387)	(0.353)	(0.410)	(0.467)	(0.396)	(0.488)
	[0.195]	[0.184]	[0.189]	[0.481]	[0.495]	[0.547]	[0.424]	[0.431]	[0.463]	[0.500]	[0.475]	[0.595]
Mon		0.014	0.032		-0.025	-0.033		-0.012	-0.007		-0.022	-0.029
		(0.203)	(0.207)		(0.215)	(0.226)		(0.208)	(0.215)		(0.213)	(0.224)
		[0.176]	[0.183]		[0.164]	[0.174]		[0.168]	[0.179]		[0.162]	[0.173]
Tues		-0.516	-0.493		-0.531	-0.533		-0.526	-0.517		-0.530	-0.530
		(0.198)	(0.204)		(0.208)	(0.220)		(0.202)	(0.210)		(0.207)	(0.218)
		[0.171]	[0.179]		[0.173]	[0.182]		[0.170]	[0.177]		[0.171]	[0.179]
Wed		-0.555	-0.539		-0.566	-0.576		-0.563	-0.561		-0.566	-0.573
		(0.202)	(0.206)		(0.213)	(0.222)		(0.207)	(0.212)		(0.211)	(0.220)
		[0.164]	[0.172]		[0.170]	[0.178]		[0.166]	[0.173]		[0.170]	[0.182]
Thurs		0.082	0.095		0.109	0.118		0.100	0.108		0.107	0.116
		(0.198)	(0.201)		(0.209)	(0.216)		(0.203)	(0.207)		(0.207)	(0.214)
		[0.170]	[0.171]		[0.178]	[0.178]		[0.171]	[0.171]		[0.173]	[0.175]
Cold			-0.062			0.068			0.015			0.059
			(0.134)			(0.173)			(0.155)			(0.167)
			[0.156]			[0.163]			[0.161]			[0.192]
Rainy			0.067			0.072			0.070			0.072
			(0.177)			(0.190)			(0.182)			(0.188)
			[0.159]			[0.160]			[0.158]			[0.160]
2SLS Chi-squared (dof=1)							0.08	.77	.90	0.41	1.47	1.29

Notes: Standard errors in parentheses were calculated using the usual asymptotic approximation for OLS and 2SLS or IV. Standard errors in square brackets were calculated using the formulas in the appendix. Chi-square statistics are for the conventional instrument-error over-identification test with one degree of freedom.

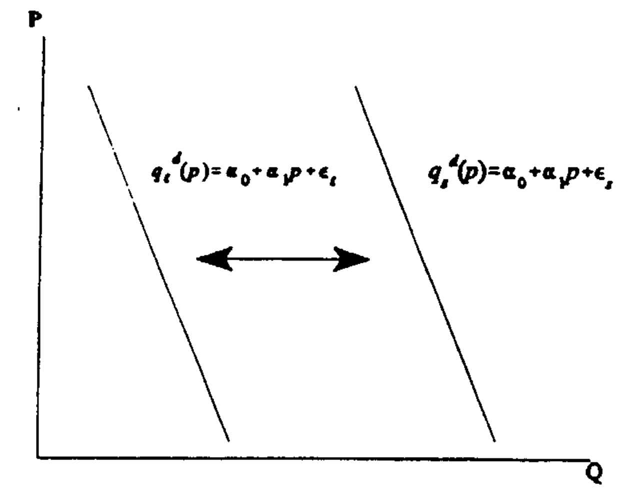

Figure 1a. Parallel shifts of a linear demand curve.

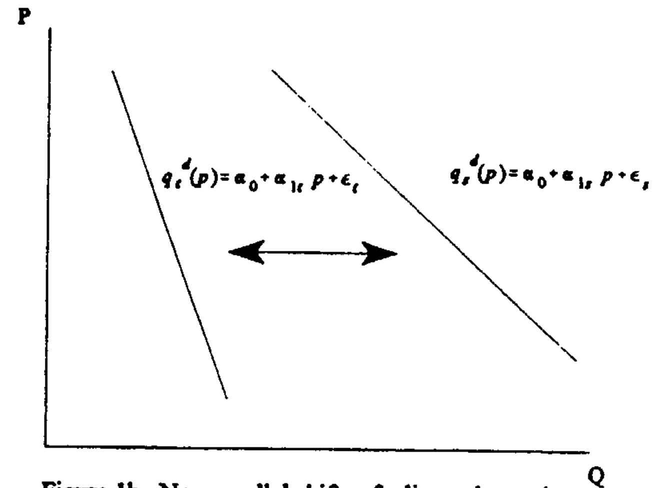

Figure 1b. Non-parallel shifts of a linear demand curve

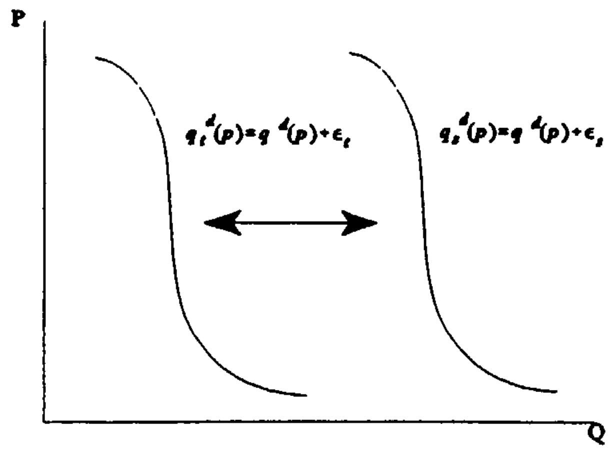

Figure 1c. Parallel shifts of a non-linear demand curve. Demand shifts in markets indexed by t and s.

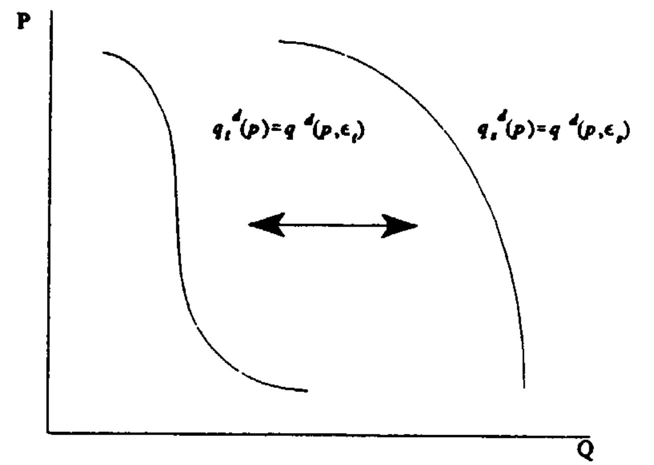

Figure 1d. Non-parallel shifts of a non-linear demand curve.

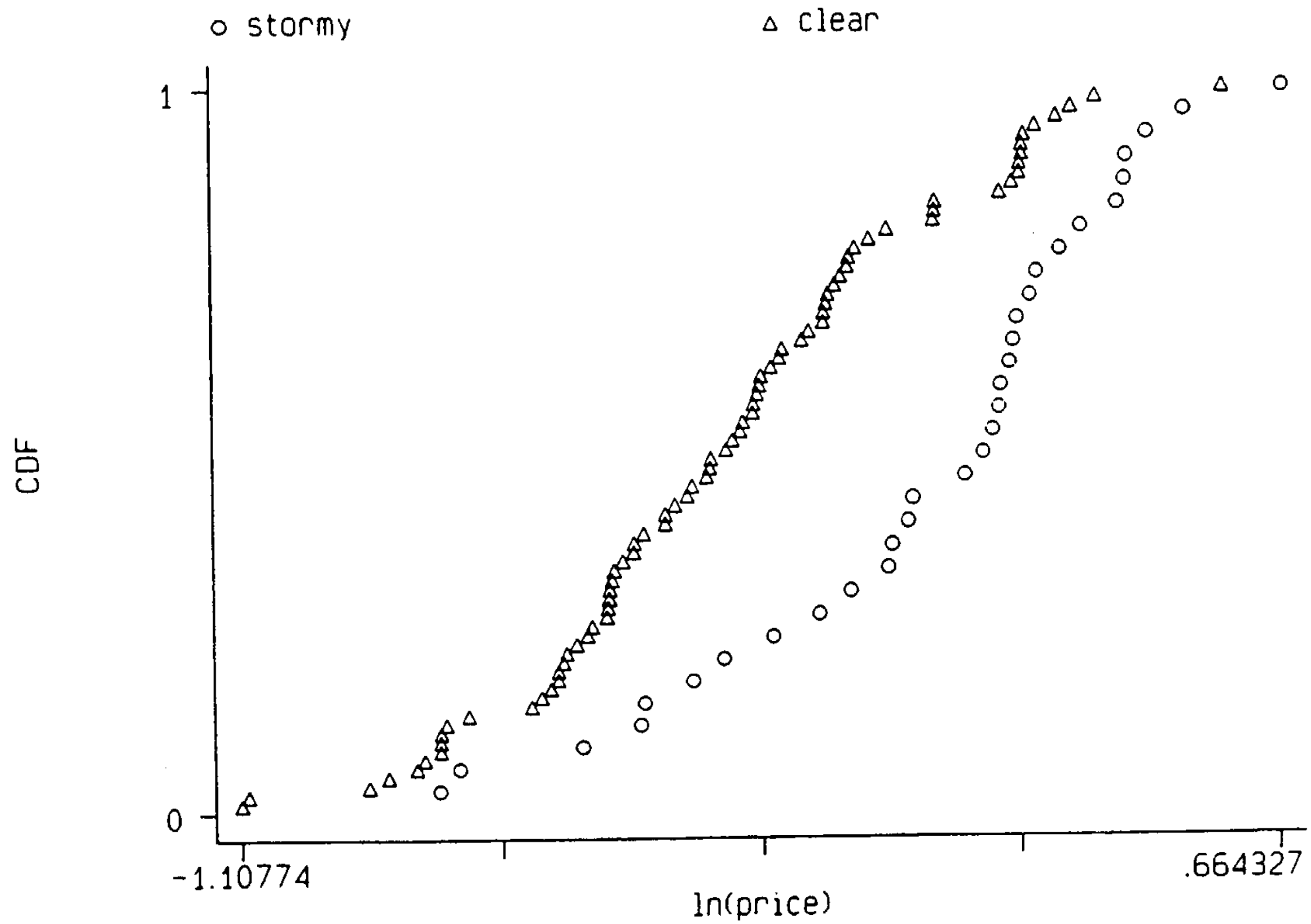

Figure 2: Weather CDFs

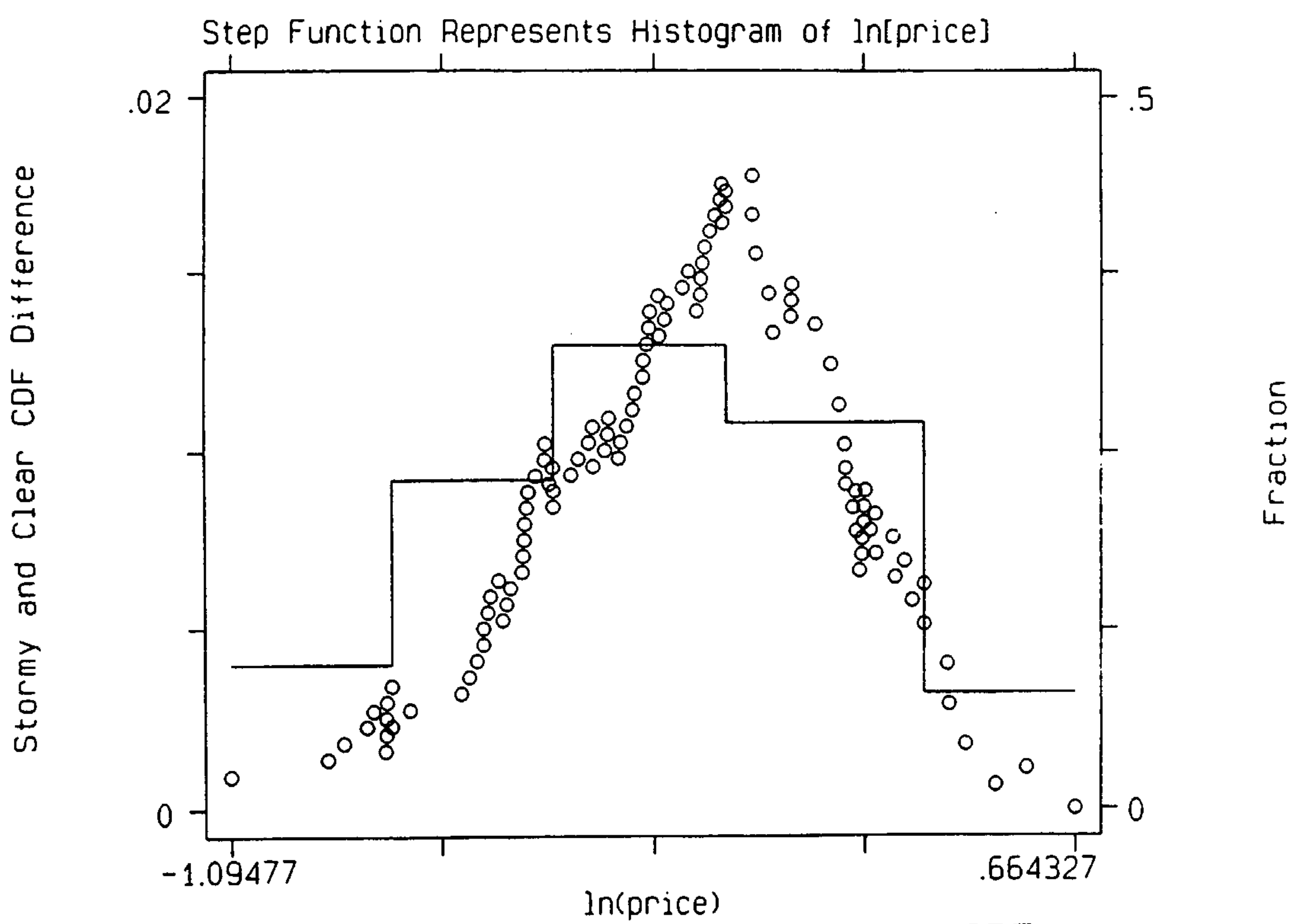

Figure 3: Difference in Weather CDFs

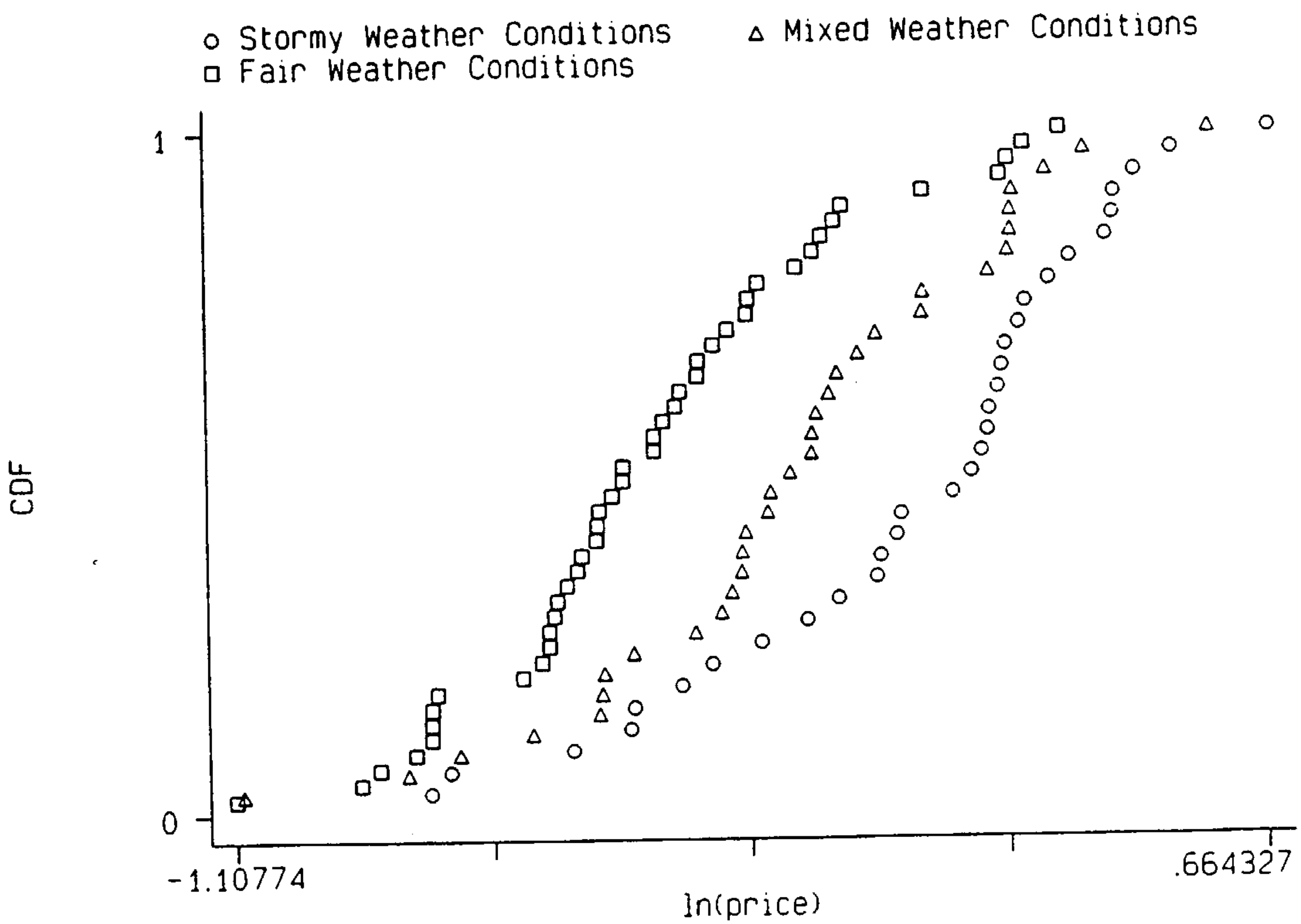

Figure 4: Weather CDFs

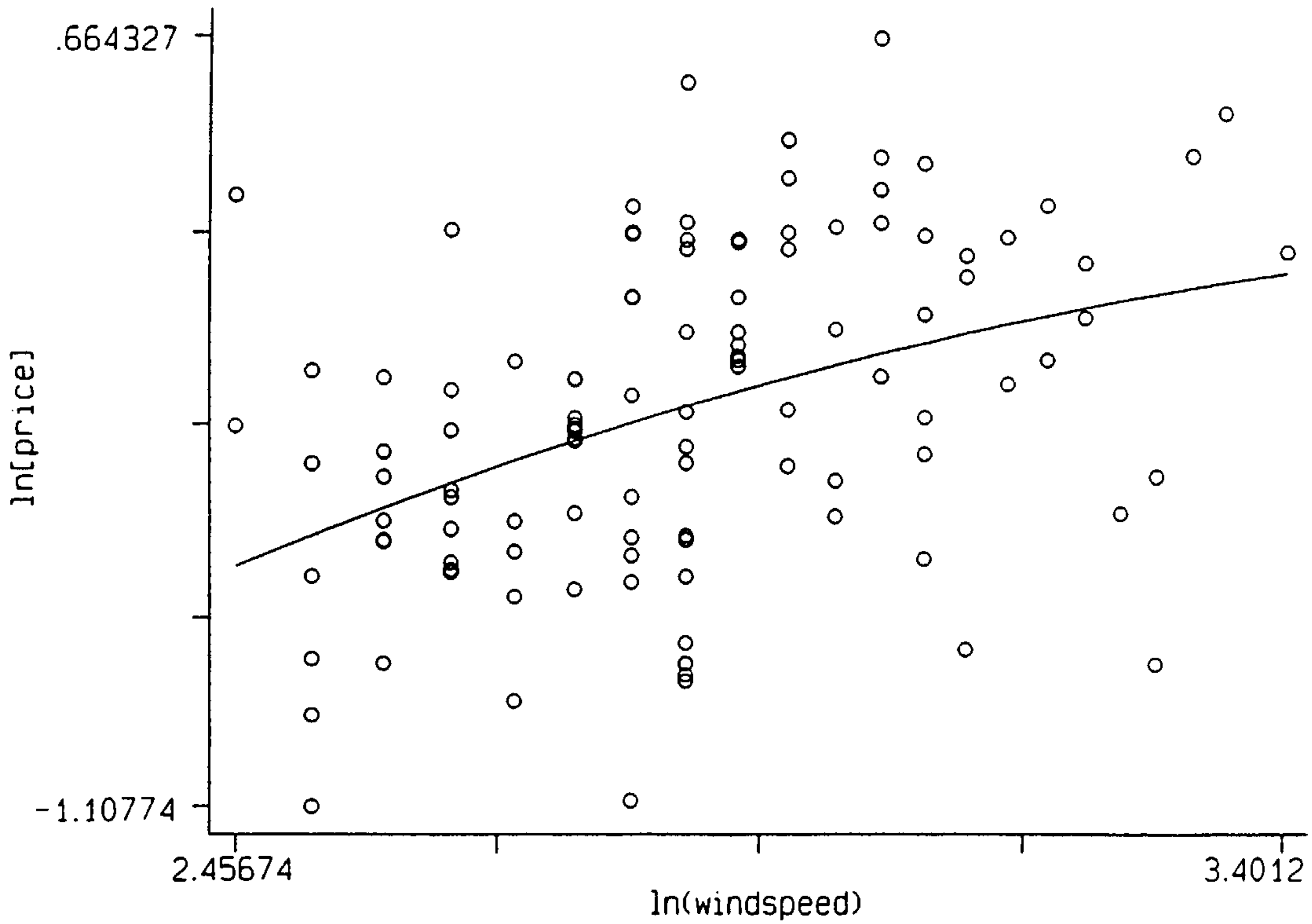

Figure 5: Price as a Function of Windspeed

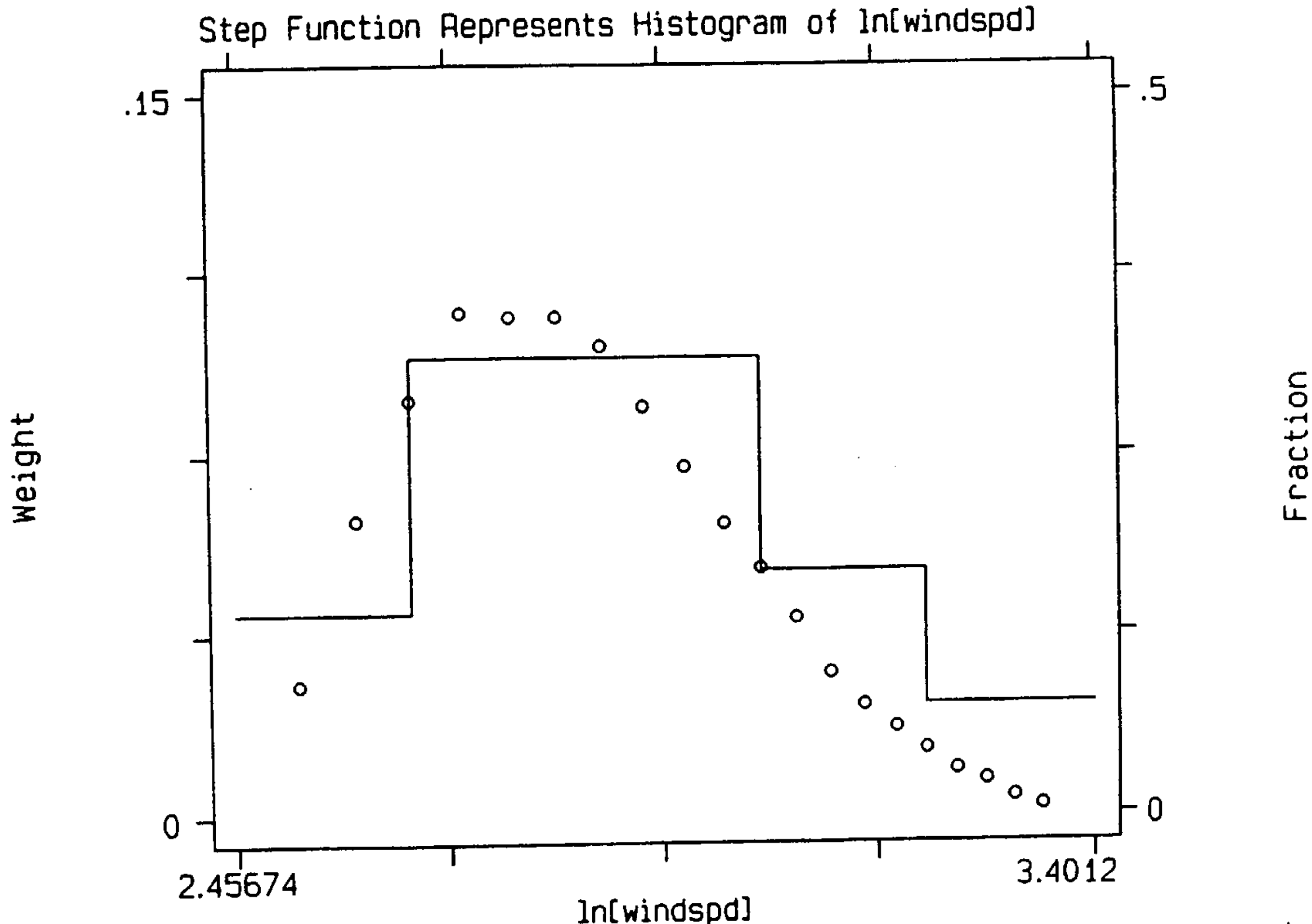

Figure 6: Weighting Function for a Continuous Instrument