Instrument-Free Demand Estimation*

Alexander MacKay
Harvard University†

Nathan H. Miller
Georgetown University‡

First Draft: March 20, 2017
This Draft: June 26, 2018

Abstract

We consider the identification and estimation of demand systems in models of imperfect competition. Under the usual assumption of profit maximization, the bias that arises from price endogeneity can be resolved without the use of instruments. In many standard demand systems, we show that the biased coefficient from an ordinary least squares regression of (transformed) quantity on price can be expressed as function of the structural demand parameters. With a covariance restriction on unobservable shocks, these parameters can be identified. Further, it can be possible to place bounds on the structural parameters without imposing a covariance restriction. We illustrate the methodology with applications to the cement industry and the airline industry.

*We thank Chuck Romeo, Gloria Sheu, Karl Schurter, Matthew Weinberg, Nathan Wilson, and seminar participants at Harvard Business School, MIT, and the University of Maryland for helpful comments.

†Harvard University, Harvard Business School. Email: amackay@hbs.edu.

‡Georgetown University, McDonough School of Business. Email: nathan.miller@georgetown.edu

1 Introduction

A central challenge of demand estimation is price endogeneity. If prices reflect demand shocks that are not observed by the econometrician, then ordinary least squares regression (OLS) does not recover the casual demand curve (Working, 1927). In this paper, we reconsider whether exogenous variation in prices is necessary to recover causal demand parameters. We show that the supply-side assumptions already maintained in many structural models dictate how prices respond to demand shocks. By leveraging these assumptions in estimation, it is possible to correct for endogeneity bias without exogenous variation in prices. The consistent estimation of empirical models has been a major focus of research in industrial organization and has, thus far, relied heavily on instruments (e.g., Berry et al. (1995); Bresnahan (1996); Hausman (1996); Berry and Haile (2014)).

Our methodology begins with an analysis of equilibrium variation in prices and (possibly transformed) quantities. We show that, with many standard empirical models of imperfect competition, the bias in OLS estimates is a function of data and demand parameters. Thus, OLS estimates are informative, as they capture a blend of the demand curve and the endogenous response by firms. The supply-side assumptions may be used to construct bounds on the structural parameters and, with the addition of a surprisingly weak assumption, achieve point identification. The methodology essentially uses economic theory as a substitute for exogenous variation in prices, allowing for consistent estimates of structural parameters without the use of instruments.

Consider a general case in which price is determined by additively separable markup and marginal cost terms, and demand takes a semi-linear form that nests the discrete-choice models common in empirical research (Berry, 1994). In this setting, the OLS bias can be decomposed into two components: (i) the covariance between demand shocks and markups and (ii) the covariance between demand shocks and marginal costs. Using the supply-side model, the first component of bias can be recovered from the data. Therefore, the surprisingly weak assumption needed for point identification relates to the covariance between the unobserved shocks to demand and marginal costs. If the econometrician has prior knowledge of this covariance, then typically the price parameter is identified.

We first develop intuition using a model of a monopolist with constant marginal costs and linear demand (Section 2). Equilibrium variation in prices and quantities (p and q) is generated by uncorrelated demand and cost shocks (ξ and η) that are unobservable to the econometrician. We prove that consistent estimate of the price parameter, β, is given by

$$\hat{\beta} = -\sqrt{\left(\hat{\beta}^{OLS}\right)^2 + \frac{Cov(\hat{\xi}^{OLS}, q)}{Var(p)}}$$

where $\hat{\beta}^{OLS}$ is the price coefficient from an OLS regression of quantities on prices, and $\hat{\xi}^{OLS}$

is a vector of the OLS residuals. The information provided by OLS regression is sufficient for the consistent identification of the structural parameters. This holds whether variation arises predominately from demand shocks or from supply shocks—economic theory allows for identification of the structural parameter even amidst a cloud of price-quantity pairs.

We obtain our baseline results (Section 3) under two common assumptions about demand and supply. We assume that demand is semi-linear in prices after a known transformation. This assumption nests many differentiated-products demand systems, including the random coefficients logit (e.g., Berry et al. (1995)). On the supply side, we begin by assuming that firms compete in prices à la Nash and have constant marginal costs. Our core identification result is that the price parameter, β, solves a quadratic equation in which the coefficients are functions of the data and the covariance between unobserved shocks to demand and marginal costs, $Cov(\xi, \eta)$. We provide a sufficient condition under which β is the lower root of the quadratic; if the condition holds then knowledge of this covariance point identifies β. We then derive a consistent three-stage estimator from the quadratic formula. The estimator is constructed from the OLS coefficients and residuals, and, based on Monte Carlo experiments, performs well in small samples.

Our three-stage estimator is developed under the assumption of *uncorrelatedness*: $Cov(\xi, \eta) = 0$. This restriction could, alternatively, be used to construct a method-of-moments estimator, which obtains identical estimates with greater computational burden. Assuming orthogonality between supply and demand shocks is not uncommon in empirical work, but the implications for identification in models of imperfect competition have not previously been formalized.[1] In other contexts, the use of covariance restrictions has been explored since early Cowles Foundation research (Koopmans et al., 1950), as we describe later.

Even without exact knowledge of $Cov(\xi, \eta)$, supply-side restrictions can be used to place bounds on β. First, weaker assumptions about $Cov(\xi, \eta)$ that are motivated by the economic environment can be used to construct bounds on the causal parameters. For example, it may be reasonable to assume that there is positive correlation between unobserved shocks to supply and demand, in which case an upper bound on β is obtained. Second, certain values of β may be ruled out without any prior knowledge of $Cov(\xi, \eta)$. We show how to construct these *prior-free* bounds, which arise when the parameter values do not rationalize the data given the assumptions of the model.

In Section 4, we relax the supply-side assumptions used to develop our baseline results. We first prove that identification of β is preserved with non-constant marginal costs if the non-constant portion can be brought into the model and estimated. We then consider multi-product firms, which is a straightforward extension of the single-product case used to develop notation earlier. Finally, we show that our approach is not dependent on the precise nature of

[1]See Thomadsen (2005), Cho et al. (2018), and Li et al. (2018) for examples in industrial organization. Thomadsen (2005) assumes no unobserved demand shocks, and Cho et al. (2018) assume no unobserved cost shocks; both implicitly invoke uncorrelatedness.

the competitive game. Instead, it relies on the general property that prices can be structurally decomposed into additively separable marginal costs and markup terms. Our identification result and three-stage estimator are easily adapted to other models of competition, including Cournot and consistent conjectures.

We provide two empirical applications in Section 5. The first examines the cement industry using the model and data of Fowlie et al. (2016) ["FRR"], extending the approach to Cournot competition. In this setting, the institutional details allow for an assessment of the uncorrelatedness assumption. Unobserved demand variation reflects local construction activity, whereas marginal cost variation is due to capacity utilization and coal prices. After incorporating capacity constraints into the model, uncorrelatedness is a reasonable assumption if local construction activity is orthogonal to coal prices. There is a theoretical basis for such an identifying assumption: if coal suppliers have limited market power and roughly constant (realized) marginal costs, then coal prices should not respond much to construction demand. Indeed, this logic motivates the use of coal prices as an instrument in FRR. Not surprisingly, a three-stage estimator obtains results similar to two-stage least squares using the FRR instruments. If capacity constraints are *not* incorporated into the model, then we expect that demand shocks drive up marginal costs via the capacity constraints, leading to a positive correlation in unobserved shocks. We show that an alternative assumption of $Cov(\xi, \eta) \geq 0$ is sufficient to place an upper bound on the price parameter that is roughly 50 percent more negative than the OLS estimate.

The second empirical application examines the airline industry using the model and data of Aguirregabiria and Ho (2012) ["AH"]. The nested logit demand system in the application has a second endogenous regressor, corresponding to the nesting parameter σ. We show how to incorporate additional restrictions to identify such parameters. Natural candidates include instruments and covariance restrictions that are generalized from the uncorrelatedness assumption. For example, if shocks are uncorrelated at the product level, it may be reasonable to assume that mean shocks are uncorrelated when aggregated by product group. Such supplemental moments are sufficient to point identify the demand system, and we show that different specifications for our three-stage estimator all move the parameter estimates in the expected direction relative to OLS. We then consider set identification under weaker assumptions. We construct prior-free bounds and bounds under the assumption of (weakly) positive correlation in shocks. In our application, we are able to rule out values of σ less than 0.599 for any value of β, as these lower values cannot generate positive correlation in both product-level and product-group-level shocks. We obtain an upper bound on β of -0.067 across all values of σ. Together, the three sets of bounds provide the identified set for (β, σ).

Section 6 provides two discussions that help frame the methodology we introduce. First, we argue that an understanding of institutional details can allow for an assessment of uncorrelatedness even though the structural error terms are (by definition) unobserved. Indeed, sometimes the institutional details will suggest that uncorrelatedness is unreasonable. Products with greater unobserved quality might be more expensive to produce, demand shocks could raise or lower marginal costs (e.g., due to capacity constraints), or firms might invest to lower the costs of their best-selling products. These cases are problematic unless the confounding variation can be absorbed by control variables or fixed effects. Second, we relate uncorrelatedness to the instrumental variables approach. The most obvious similarity is that both approaches rely on orthogonality conditions that are not verifiable empirically but can be assessed with institutional details. This connection is especially clear with the so-called "Hausman" instruments—prices of the same good in other markets—for which consistency requires orthogonality among demand shocks across markets. However, the assumptions embedded by the two approaches are not generally nested: uncorrelatedness does not require any source of exogenous variation but does require a correctly-specified supply-side model.

Our research builds on several strands of literature in economics. Early research at the Cowles Foundation (Koopmans et al., 1950) examines the identifying power of covariance restrictions in linear systems of equations, and a number of articles pursued this agenda in subsequent years (e.g., Fisher (1963, 1965); Wegge (1965); Rothenberg (1971); Hausman and Taylor (1983)). The extension to semi-parametric models is provided in Matzkin (2016) and Chiappori et al. (2017)), but market power is not considered. To help develop the connections between methodologies, we provide a new identification proof for perfect competition that uses the techniques developed herein (Appendix A).

A parallel literature examines the identification of supply and demand models using maximum likelihood techniques, under the assumptions that the distributions of demand and cost shocks are known to the econometrician and independent. Leamer (1981) provides conditions under which the price parameter can be bounded using only the endogenous variation in prices and quantities. Feenstra (1994) extends the methodology to estimate a model of monopolistic competition.[2] Using Bayesian techniques, Yang et al. (2003) further extends the approach to oligopoly. Published comments on this article point out that it is unclear how to write a coherent likelihood function for oligopoly games because multiple equilibria can exist (Bajari, 2003; Berry, 2003). By contrast, our approach does not require a likelihood function and provides consistent estimates in the presence of multiple equilibria. Further, it allows the econometrician to relax distributional assumptions. These advantages may make our approach relatively more palatable for oligopoly models.[3]

Price endogeneity has been a major focus of modern empirical and econometric research in industrial organization. Typically, the challenge is cast as a problem of finding valid instruments. Many possibilities have been developed, including the attributes of competing products

[2] Leamer attributes an early version of his results to Schultz (1928). Broda and Weinstein (2006) and Hottman et al. (2016) have extended this approach in models of international trade.

[3] For discussions and extensions of the Yang et al. (2003) approach in the marketing literature, see Rossi et al. (2005), Dotson and Allenby (2010), and Otter et al. (2011). At least one seminal article in industrial organization, Bresnahan (1987), estimates an oligopoly model of supply and demand with maximum likelihood under the assumption of independent shocks.

(Berry et al., 1995; Gandhi and Houde, 2015), the prices of the same good in other markets (e.g., Hausman (1996); Nevo (2001); Crawford and Yurukoglu (2012)), or shifts in the equilibrium concept (e.g., Porter (1983); Miller and Weinberg (2017)).[4] When valid instruments are available, the estimation techniques presented here may be used to construct overidentifying restrictions and test the model.

2 A Motivating Example: Monopoly Pricing

We introduce the supply-side identification approach with a motivating example of monopoly pricing, in the spirit of Rosse (1970). In each market $t = 1, \ldots, T$, the monopolist faces a downward-sloping linear demand schedule, $q_t = \alpha + \beta p_t + \xi_t$, where q_t and p_t denote quantity and price, respectively, $\beta < 0$ is the price parameter, and ξ_t is mean-zero stochastic demand shock. Marginal cost is given by the function $c_t = \gamma + \eta_t$, where γ is some constant and η_t is a mean-zero stochastic cost shock. Prices are set to maximize profit. The econometrician observes vectors of prices, $p = [p_1, p_2, \ldots, p_T]'$, and quantities, $q = [q_1, q_2, \ldots, q_T]'$. The markets can be conceptualized as geographically or temporally distinct.

An OLS regression of q on p obtains a biased estimate of β if the monopolist's price reflects the unobservable demand shock, as is the case here given profit maximization. Formally,

$$\hat{\beta}^{OLS} = \frac{Cov(p,q)}{Var(p)} \quad \xrightarrow{p} \quad \beta + \frac{Cov(\xi,p)}{Var(p)} \tag{1}$$

The monopolist's profit-maximization conditions are such that price is equal to marginal cost plus a markup term: $p_t = \gamma + \eta_t - \left(\frac{dq}{dp}\right)^{-1} q_t$. Thus, the numerator of the OLS bias can be decomposed into the covariance between demand shocks and markups and the covariance between demand shocks and marginal cost shocks. This leads to our first theoretical result, which we obtain under the uncorrelatedness assumption that $Cov(\xi, \eta) = 0$:

Proposition 1. *Let the OLS estimates of (α, β) be $(\hat{\alpha}^{OLS}, \hat{\beta}^{OLS})$ with probability limits $(\alpha^{OLS}, \beta^{OLS})$, and denote the residuals at the limiting values as $\xi_t^{OLS} = q_t - \alpha^{OLS} - \beta^{OLS} p_t$. When demand shocks and cost shocks are uncorrelated, the probability limit of the OLS estimate can be expressed as a function of the true price parameter, the residuals from the OLS regression, prices, and quantities:*

$$\beta^{OLS} \equiv plim\left(\hat{\beta}^{OLS}\right) = \beta - \frac{1}{\beta + \frac{Cov(p,q)}{Var(p)}} \frac{Cov(\xi^{OLS}, q)}{Var(p)} \tag{2}$$

Proof: We provide the proofs in this section for illustrative purposes; most subse-

[4] Byrne et al. (2016) proposes a novel set of instruments that leverage the structure of a discrete choice demand model with differentiated-products price competition. Nevo and Wolfram (2002) explores whether covariance restrictions can bound parameters (see footnote 41 of that article).

quent proofs are confined to the appendix. Reformulate equation (1) as follows:

$$\beta^{OLS} = \beta + \frac{Cov(\xi, \eta - \frac{1}{\beta}q)}{Var(p)} = \beta - \frac{1}{\beta}\frac{Cov(\xi, q)}{Var(p)}$$

The first equality holds due to the first-order condition $p = \gamma + \eta_t - \frac{1}{\beta}q$. The second equality holds due to the uncorrelatedness assumption. As $\alpha + \beta p + \xi = \alpha^{OLS} + \beta^{OLS}p + \xi^{OLS}$, we have

$$\begin{aligned} Cov(\xi, q) &= Cov(\xi^{OLS} - (\beta - \beta^{OLS})p, q) \\ &= Cov(\xi^{OLS}, q) - (\beta - \beta^{OLS})Cov(p, q) \\ &= Cov(\xi^{OLS}, q) - \frac{1}{\beta}\frac{Cov(\xi, q)}{Var(p)}Cov(p, q) \end{aligned}$$

Collecting terms and rearranging implies

$$\frac{1}{\beta}Cov(\xi, q) = \frac{1}{\beta + \frac{Cov(p,q)}{Var(p)}}Cov(\xi^{OLS}, q)$$

Plugging into the reformulation of equation (1) obtains the proposition. QED.

The proposition makes clear that, among the objects that characterize β^{OLS}, only β does not have a well understood sample analog. Further, as β^{OLS} can be estimated consistently, the proposition suggests the possibility that β can be recovered from the data. Indeed, a closer inspection of equation (2) reveals that β solves a quadratic equation:

Proposition 2. *When demand shocks and marginal cost shocks are uncorrelated, β is point identified as the lower root of the quadratic equation*

$$\beta^2 + \beta\left(\frac{Cov(p, q)}{Var(p)} - \beta^{OLS}\right) + \left(-\frac{Cov(\xi^{OLS}, q)}{Var(p)} - \frac{Cov(p, q)}{Var(p)}\beta^{OLS}\right) = 0 \tag{3}$$

and a consistent estimate of β is given by

$$\hat{\beta}^{\text{3-Stage}} = -\sqrt{\left(\hat{\beta}^{OLS}\right)^2 + \frac{Cov(\hat{\xi}^{OLS}, q)}{Var(p)}} \tag{4}$$

Proof: The quadratic equation is obtained as a re-expression of equation (2). An application of the quadratic formula provides the following roots:

$$\frac{-\left(\frac{Cov(p,q)}{Var(p)} - \beta^{OLS}\right) \pm \sqrt{\left(\frac{Cov(p,q)}{Var(p)} - \beta^{OLS}\right)^2 + 4\left(\frac{Cov(\xi^{OLS},q)}{Var(p)} + \frac{Cov(p,q)}{Var(p)}\beta^{OLS}\right)}}{2}.$$

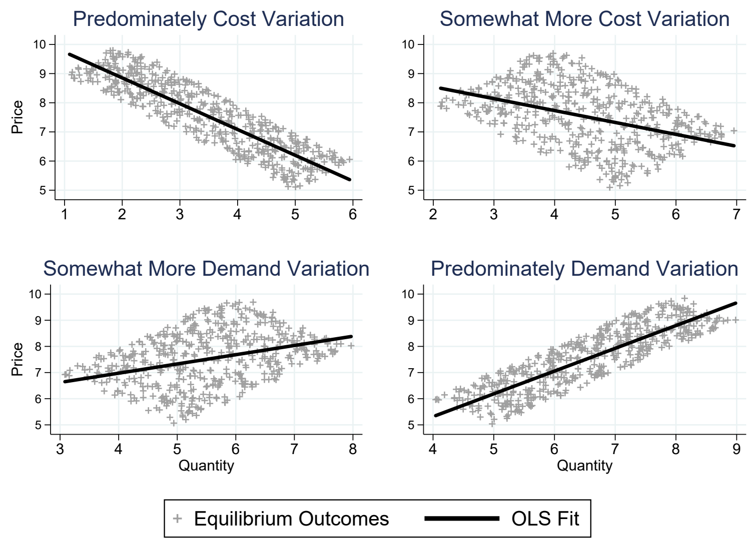

Figure 1: Price and Quantity in the Monopoly Model
Notes: Figure displays equilibrium prices and quantities under four different specifications for the distribution of unobserved shocks to demand and marginal costs. The line in each figure indicates the slope obtained by OLS regression.

In the univariate case, $\frac{Cov(p,q)}{Var(p)} = \beta^{OLS}$, which cancels out terms and obtains the probability limit analog of equation (4). It is easily verified that $\left(\beta^{OLS}\right)^2 + \frac{Cov(\xi^{OLS},q)}{Var(p)} > 0$ so both roots are real numbers. The upper root is positive, so β is point identified as the lower root. Equation (4) provides the empirical analog to the lower root. As the sample estimates of covariance terms are consistent for the limits, it provides a consistent estimate of β.

The first part of the proposition states that β solves a quadratic equation. There are two real roots, but only one is negative, so point identification is achieved. Further, an adjustment to the OLS estimator is sufficient to correct for bias. We label the estimator $\hat{\beta}^{\textit{3-Stage}}$ for reasons that become evident with the more general treatment later in the paper.

An additional simplification is available. Because $\xi_t^{OLS} = q_t - a^{OLS} - b^{OLS} p_t$, we have $\frac{Cov(\xi_t^{OLS},q)}{Var(p)} = \frac{Var(q)}{Var(p)} - \beta^{OLS} \frac{Cov(p,q)}{Var(p)}$. Plugging into equation (4) obtains the following corollary:

Corollary 1. $\hat{\beta}^{\text{3-Stage}} = -\sqrt{\frac{Var(q)}{Var(p)}}$.

In the monopoly model, the price parameter is identified from the relative variation in prices and quantities. To build intuition about this approach, we recast the monopoly problem in terms of supply and demand in Appendix A.1, and derive the estimator building on Hayashi's (2000) textbook treatment of bias with simultaneous equations.

Consider the following numerical example. Let demand be given by $q_t = 10 - p_t + \xi_t$ and let marginal cost be $c_t = \eta_t$, so that $(\alpha, \beta, \gamma) = (10, -1, 0)$. Let the demand and cost shocks

Table 1: Numerical Illustration for the Monopoly Model

	(1)	(2)	(3)	(4)
$\hat{\beta}^{OLS}$	-0.89	-0.42	0.36	0.88
$Var(q)$	1.47	1.11	1.08	1.38
$Var(p)$	1.45	1.09	1.06	1.37
$Cov(\hat{\xi}^{OLS}, q)$	0.31	0.92	0.94	0.32
$Cov(\hat{\xi}^{OLS}, q)/Var(p)$	0.21	0.85	0.89	0.24
$\hat{\beta}^{\textit{3-Stage}}$	-1.004	-1.009	-1.009	-1.004

Notes: Based on numerically generated data that conform to the motivating example of monopoly pricing. The demand curve is $q_t = 10-p_t+\xi_t$ and marginal costs are $c_t = \eta_t$, so that $(\beta_0, \beta, \gamma_0) = (10, -1, 0)$. In column (1), $\xi \sim U(0, 2)$ and $\eta \sim U(0, 8)$. In column (2), $\xi \sim U(0, 4)$ and $\eta \sim U(0, 6)$. In column (3), $\xi \sim U(0, 6)$ and $\eta \sim U(0, 4)$. In column (4), $\xi \sim U(0, 8)$ and $\eta \sim U(0, 2)$. Thus, the support of the cost shocks are largest (smallest) relative to the support of the demand shocks in the left-most (right-most) column.

have independent uniform distributions. The monopolist sets price to maximize profit. As is well known, if both cost and demand variation is present then equilibrium outcomes provide a "cloud" of data points that do not necessarily correspond to the demand curve. To illustrate this, we consider four cases with varying degrees of cost and demand variation. In case (1), $\xi \sim U(0, 2)$ and $\eta \sim U(0, 8)$. In case (2), $\xi \sim U(0, 4)$ and $\eta \sim U(0, 6)$. In case (3), $\xi \sim U(0, 6)$ and $\eta \sim U(0, 4)$. In case (4), $\xi \sim U(0, 8)$ and $\eta \sim U(0, 2)$. We randomly take 1,000 draws for each case and calculate the equilibrium prices and quantities.

The data are plotted in Figure 1 along with OLS fits. The experiment represents the classic identification problem of demand estimation: the empirical relationship between equilibrium prices and quantities can be quite misleading about the slope of the demand function. However, Proposition 2 and Corollary 1 state that the structure of the model together with the OLS estimates allow for consistent estimates of the price parameter. Table 1 provides the required empirical objects. The OLS estimates, $\hat{\beta}^{OLS}$, are negative when the cost shocks are relatively more important and positive when the demand shocks are relatively more important, as also is revealed in the scatterplots. By contrast, $\frac{Cov(\hat{\xi}^{OLS},q)}{Var(p)}$ is larger if the cost and demand shocks have relatively more similar support. Incorporating this adjustment term following Proposition 1 yields estimates, $\hat{\beta}^{\textit{3-Stage}}$, that are nearly equal to the population value of -1.00. Note also that the variance of price and quantity are similar in each column, consistent with Corollary 1 given the data generating process.

3 Methodology: Bounds and Three-Stage Estimation

We present our main results through the lens of differentiated-products Bertrand competition. We provide identification conditions, show how bounds can be constructed without strong covariance restrictions, and then consider estimation under uncorrelatedness via the three-stage approach and the method of moments. We illustrate small-sample properties with a numerical simulation. For extensions and a discussion, see Sections 4 and 6.

3.1 Data Generating Process

Let there be $j = 1, 2, \ldots, J$ products in each of $t = 1, 2, \ldots, T$ markets, subject to downward-sloping demands. The econometrician observes vectors of prices, $p_t = [p_{1t}, p_{2t}, \ldots, p_{Jt}]'$, and quantities, $q = [q_{1t}, q_{2t}, \ldots, p_{Jt}]'$, corresponding to each market t, as well as a matrix of covariates $X_t = [x_{1t}\ x_{2t}\ \ldots\ x_{Jt}]$. The covariates are orthogonal to a pair of demand and marginal cost shocks (i.e,. $E[X\xi] = E[X\eta] = 0$) that are common knowledge among firms but unobserved by the econometrician.[5] We make the following assumptions on demand and supply:

Assumption 1 (Demand): *The demand schedule for each product is determined by the following semi-linear form:*

$$h_{jt} \equiv h(q_{jt}, w_{jt}; \sigma) = \beta p_{jt} + x'_{jt}\alpha + \xi_{jt} \tag{5}$$

where (i) $\frac{\partial h_{jt}}{\partial q_{jt}} > 0$*, (ii)* w_{jt} *is a vector of observables and* σ *is a parameter vector, and (iii) the total derivatives of* $h(\cdot)$ *with respect to* q *exist as functions of the data and* σ*.*

> *Example:* For the logit demand system, $h(q_{jt}; w_{jt}, \sigma) \equiv \ln(s_{jt}/w_{jt})$, where quantities are in shares ($q_{jt} = s_{jt}$) and w_{jt} is the share of the outside good (s_{0t}). There are no additional parameters in σ.

The demand assumption restricts attention to systems for which, after a transformation of quantities using observables (w_{jt}) and nonlinear parameters (σ), there is additive separability in prices, covariates, and the demand shock. The vector w_{jt} can be conceptualized as including the price and non-price characteristics of products, in particular those of other products that affect the demand of product j. Often, only a few observables are necessary to construct the transformation, as in the logit discrete choice example above, where the share of the outside good is a sufficient statistic to capture demand for other products.

Other typical demand systems also fit into this framework, including models with additional endogenous covariates. For the nested logit demand model, w_{jt} consists of two elements: the outside share (s_{0t}) and the within-group share ($\bar{s}_{j|g}$). One nonlinear parameter (σ) is needed for the transformation: $h(s_{jt}; w_{jt}, \sigma) \equiv \ln s_{jt} - \ln s_{0t} - \sigma \ln \bar{s}_{j|g,t}$. For the more flexible random coefficients logit demand system, $h(q_{jt}; w_{jt}, \sigma)$ can be defined as the mean utility and calculated using the contraction mapping of Berry et al. (1995) for any candidate σ vector. The demand assumption also nests monopolistic competition with linear demands (e.g., as in the motivating example). We derive these connections in some detail in Appendix B.

[5] We make the usual assumption that prices and covariates are linearly independent to allow for OLS estimation.

The third condition on the demand system allows us to complete an identification proof by constructing the first-order conditions implied by the demand system and the supply-side assumptions, which we introduce next. The total derivative is given by

$$\frac{dh_{jt}}{dq_{jt}} = \frac{\partial h_{jt}}{\partial q_{jt}} + \frac{\partial h_{jt}}{\partial w_{jt}} \cdot \frac{dw_{jt}}{dq_{jt}}$$

When an outcome variable is used in w_{jt} to construct the transformation, such as in the discrete choice demand systems mentioned above, it may be the case that $\frac{dw_{jt}}{dq_{jt}}$ depends upon what is held fixed under the competitive assumptions (e.g. the prices of other firms at the Bertrand-Nash equilibrium.). This should be accounted for when constructing the derivatives.

> *Example:* For the logit demand system at a Bertrand-Nash equilibrium, $\frac{\partial h_{jt}}{\partial q_{jt}} = \frac{1}{s_{jt}}$, $\frac{\partial h_{jt}}{\partial w_{jt}} = -\frac{1}{s_{0t}}$, and $\frac{dw_{jt}}{dq_{jt}} = \frac{ds_{0t}/dp_{jt}}{ds_{jt}/dp_{jt}} = -\frac{s_{0t}}{1-s_{jt}}$. Thus we obtain $\frac{dh_{jt}}{dq_{jt}} = \frac{1}{s_{jt}(1-s_{jt})}$. These derivatives are calculated holding the prices of other products fixed.

Assumption 2 (Supply): *Each firm sells a single product and sets its price to maximize profit in each market. The firm takes the prices of other firms as given, knows the demand schedule in equation* (5)*, and has a linear constant marginal cost schedule given by*

$$c_{jt} = x_{jt}'\gamma + \eta_{jt}. \tag{6}$$

Under assumptions 1 and 2, there is a unique mapping from the data and parameters to the structural error terms (ξ, η). The supply-side assumption is strong but allows for a base set of identification results to be derived with minimal notation. In subsequent sections, we provide the additional notation necessary for models with multi-product firms, non-constant marginal costs, and Nash-Cournot competition. Note that supply and demand may depend on different covariates; this is captured when non-identical components of α and γ are equal to zero.

We further assume the existence of a Nash equilibrium in pure strategies, and that each firm satisfies the first-order condition

$$p_{jt} = c_{jt} - \frac{1}{\beta}\frac{dh_{jt}}{dq_{jt}} q_{jt}. \tag{7}$$

To obtain this expression, take the total derivative of h with respect to q_{jt}, re-arrange to obtain $\frac{dp_{jt}}{dq_{jt}} = \frac{1}{\beta}\frac{dh_{jt}}{dq_{jt}}$, and substitute into the more standard formulation of the first-order condition: $p = c - \frac{dp}{dq}q$. First-order conditions that admit multiple equilibria are unproblematic. It must be possible recover (ξ, η) from the data and parameters, but the mapping to prices from the parameters, exogenous covariates, and structural error terms need not be unique.

Our identification result relies on the markup being proportional to the reciprocal of the price parameter, which arises here due to the semi-linear demand system.[6] When this is the

[6] Thus, the semi-linear structure may not be necessary. In practice, one could start with a known first-order

case, equilibrium prices respond to the demand shock through markup adjustments, which are fully determined by β, the structure of the model, and observables. Thus, first-order conditions may be useful in analyzing the covariance of demand shocks and prices, which is proportional to the the bias of the OLS estimate. As in the monopoly example of Section 2, this provides a basis to correct the bias from OLS estimation and solve for the true price parameter. We develop this identification argument below and then derive implications for inference.

3.2 Identification

We now formalize the identification argument for β, the price parameter. We assume the parameters in σ are known to the econometrician. The linear non-price parameters (α, γ) can be recovered trivially given β and σ.[7] We start by characterizing the OLS estimate of the price parameter, which is obtained from a regression of $h(\cdot)$ on p and x. The probability limit contains the standard bias term:

$$\beta^{OLS} \equiv \frac{Cov(p^*, h)}{Var(p^*)} = \beta + \frac{Cov(p^*, \xi)}{Var(p^*)} \tag{8}$$

where $p^* = [I - x(x'x)^{-1}x']p$ is a vector of residuals from a regression of p on x. Plugging in for price on the right-hand-side of equation (8) using the first-order conditions yields

$$\beta^{OLS} = \beta - \frac{1}{\beta}\frac{Cov(\frac{dh}{dq}q, \xi)}{Var(p^*)} + \frac{Cov(\eta, \xi)}{Var(p^*)}. \tag{9}$$

We express the unobserved demand shock ξ in terms of the OLS residuals and parameters to obtain our first general result: β solves a quadratic equation in which the coefficients are determined by $Cov(\xi, \eta)$ and objects with empirical analogs.

Proposition 3. *Under assumptions 1 and 2, the probability limit of the OLS estimate can be written as a function of the true price parameter, the residuals from the OLS regression, the covariance between demand and supply shocks, prices, and quantities:*

$$\beta^{OLS} = \beta - \frac{1}{\beta + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}}\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} + \beta\frac{1}{\beta + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}}\frac{Cov(\xi, \eta)}{Var(p^*)}. \tag{10}$$

---

condition and show that it takes the form $p_{jt} = c_{jt} - \frac{1}{\beta}f_{jt}$ for some function of the data f_{jt}.

[7] An alternative interpretation is that the econometrician is considering a candidate σ and wishes to obtain corresponding estimates of (β, α, γ), as in the nested fixed-point estimation routine of Berry et al. (1995) and Nevo (2001) for the random coefficients logit demand system.

The price parameter β solves the following quadratic equation:

$$\begin{aligned} 0 &= \beta^2 \\ &+ \left(\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)} - \beta^{OLS}\right)\beta \\ &+ \left(-\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\right). \end{aligned} \tag{11}$$

Proof. See appendix.

Proposition 3 provides our core identification result. There are two main implications. First, the quadratic in equation (11) admits at most two solutions for a given value of $Cov(\xi,\eta)$. It follows immediately that, with prior knowledge of $Cov(\xi,\eta)$, the price parameter β is set identified with a maximum of two elements (points). Indeed, as we show next, conditions exist that guarantee point identification. Second, if the econometrician does not have specific knowledge of $Cov(\xi,\eta)$, it nonetheless can be possible to bound β. We consider point identification first, as the intuition behind point identification maps neatly into how to construct bounds.

Assumption 3': *The econometrician has prior knowledge of $Cov(\xi,\eta)$.*

Proposition 4. ***(Point Identification)*** *Under assumptions 1 and 2, the price parameter β is the lower root of equation* (11) *if the following condition holds:*

$$0 \leq \beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} \tag{12}$$

and, furthermore, β is the lower root of equation (11) *if and only if the following condition holds:*

$$-\frac{1}{\beta}\frac{Cov(\xi,\eta)}{Var(p^*)} \leq \frac{Cov\left(p^*, -\frac{1}{\beta}\xi\right)}{Var(p^*)} + \frac{Cov\left(p^*,\eta\right)}{Var(p^*)} \tag{13}$$

Therefore, under assumptions 1, 2 and 3', β is point identified if either of these conditions holds.

Proof. See appendix.

The first (sufficient) condition is derived as a simple application of the quadratic formula: if the constant term in the quadratic of equation (11) is negative then the upper root of the quadratic is positive and β must be the lower root. For some model specifications, the condition can be proven analytically.[8] Otherwise it can be evaluated empirically using the data and

[8]An example is a monopolist with a linear demand system. Following the logic of Corollary 1, we have

$$\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} \quad = \quad \frac{Var(q)}{Var(p)} > 0$$

assumptions 1 and 2. If the sufficient condition holds, then β is point identified with prior knowledge of $Cov(\xi, \eta)$ because all the terms in equation (11) are known or can be obtained from the data. If the condition fails, point identification of β is not guaranteed even with prior knowledge of $Cov(\xi, \eta)$, though the econometrician has reduced the identified set to two points.

The necessary and sufficient condition is more nuanced. Even with prior knowledge of $Cov(\xi, \eta)$, condition (13) contains elements that are not observed by the econometrician. Still, for some specifications, the condition can be verified analytically.[9] The condition holds under the standard intuition that prices increase both with demand and marginal cost shocks, provided that $Cov(\xi, \eta)$ is not too positive. To see this in the equation, note that the term $-\frac{1}{\beta}\xi$ is the shock to the inverse demand curve. The condition can fail if the empirical variation is driven predominately by demand shocks and the model dictates that prices decrease in the demand shock, which is possible with log-convex demand (Fabinger and Weyl, 2014).

3.3 Bounds

The model implies two complementary sets of bounds, neither of which requires exact knowledge of $Cov(\xi, \eta)$. We start by developing what we refer to as *bounds with priors*. If the econometrician has a prior over the plausible range of $Cov(\xi, \eta)$, along the lines of $m \leq Cov(\xi, \eta) \leq n$, then a posterior set for β can be constructed from the quadratic of equation (11). Each plausible $Cov(\xi, \eta)$ maps into one or two valid (i.e., negative) roots. Further, a monotonicity result that we formalize below establishes that, under either condition (12) or (13), there is a one-to-one mapping between the value of $Cov(\xi, \eta)$ and the lower root:

Lemma 1. ***(Monotonicity)*** *Under assumptions 1 and 2, a valid lower root of equation* (11) *(i.e., one that is negative) is decreasing in* $Cov(\xi, \eta)$. *The range of the function is* $(0, -\infty)$.

Proof. See appendix.

It follows immediately that a convex prior over $Cov(\xi, \eta)$ corresponds to convex posterior set. We suspect that, in practice, most priors will take the form $Cov(\xi, \eta) \geq 0$ or $Cov(\xi, \eta) \leq 0$. For example, an econometrician have reason to believe that higher quality products are more expensive to produce (yielding $Cov(\xi, \eta) \geq 0$) or that firms invest to lower the marginal costs of their best-selling products (yielding $Cov(\xi, \eta) \leq 0$). Priors of this firm generate one-sided bounds on β. Let $r(m)$ be the lower root of the quadratic evaluated at $Cov(\xi, \eta) = m$.

[9] Consider again the example of a monopoly facing a linear demand system with $Cov(\xi, \eta) = 0$. In the proof of Proposition 4, we show that the necessary and sufficient condition is equivalent to $\beta < \frac{Cov(p^*, \xi)}{Var(p^*)} - \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)}$. With linear demand, we have that $\frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} = \frac{Cov(p^*, q)}{Var(p^*)} = \beta^{OLS}$. Thus, the right-hand-side simplifies to $-\beta$ using equation (8). Because $\beta < 0$, $\beta < -\beta$ and the necessary and sufficient condition holds.

Then under either condition (12) or (13), the prior $Cov(\xi, \eta) \geq m$ produces a posterior set of $(-\infty, r(m)]$, and the prior $Cov(\xi, \eta) \leq m$ produces a posterior set of $[r(m), 0)$.[10]

We now develop what we refer to as *prior-free bounds*. Even if the econometrician has no prior about $Cov(\xi, \eta)$, certain values may be possible to rule out because they imply that the observed data are incompatible with data generating process of the model. To see why, it is helpful to represent the quadratic of equation (11) as az^2+bz+c, keeping in mind that one root is $\beta < 0$. Because $a = 1$, the quadratic forms a $\cup$-shaped parabola. If $c < 0$ then the existence of a negative root is guaranteed. However, if $c > 0$ then b must be positive and sufficiently large for a negative root to exist. By inspection of equation (11), this places restrictions on $Cov(\xi, \eta)$. We now state the result formally:

Proposition 5. ***(Prior-Free Bound)*** *Under assumptions 1 and 2, the model and data may bound $Cov(\xi, \eta)$ from below. The bound is given by*

$$Cov(\xi,\eta) > Var(p^*)\beta^{OLS} - Cov(p^*, \frac{dh}{dq}q) + 2Var(p^*)\sqrt{\left(-\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\right)}.$$

The bound exists if and only if the term inside the radical is non-negative.

Proof. See appendix.

From the monotonicity result above, we can use the excluded values of $Cov(\xi, \eta)$ from this result to rule out values of β as well. If point identification can be shown via the necessary and sufficient condition, then an prior-free upper bound for β is obtained by evaluating the lower root of equation (11) at the prior-free bound of $Cov(\xi, \eta)$.[11]

3.4 Estimation

The consistent estimation of β is possible if the conditions for point identification hold. The econometrician must have prior knowledge of $Cov(\xi, \eta)$. For the purposes of exposition, we proceed here under the uncorrelatedness assumption, $Cov(\xi, \eta) = 0$, though the mathematics extend to alternative restrictions.

Assumption 3 (Uncorrelatedness): $Cov(\xi, \eta) = 0$.

There are two natural approaches to estimation. The first is to apply the quadratic formula directly to equation (11). The second is to recast uncorrelatedness as a moment restriction of the form $E[\xi'\eta] = 0$ and use the method of moments. Of these, the first is more novel, and so we open this section with the relevant theoretical result:

[10] Nevo and Rosen (2012) develop similar bounds for estimation with imperfect instruments, defined as instruments that are less correlated with the structural error term than the endogenous regressor.

[11] Interestingly, prior-free bounds are available only if the sufficient condition for point identification (condition (12)) fails. When this occurs the term inside the radical is non-negative.

Corollary 2. ***(Three-Stage Estimator)*** *Under assumptions 1, 2, and 3, a consistent estimate of the price parameter β is given by*

$$\hat{\beta}^{\text{3-Stage}} = \frac{1}{2}\left(\hat{\beta}^{OLS} - \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)} - \sqrt{\left(\hat{\beta}^{OLS} + \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\hat{\xi}^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}}\right)$$

if either condition (12) *or condition* (13) *holds.*

The estimator is the empirical lower root of equation (11). It can be calculated in three stages: (i) regress $h(q)$ on p and x with OLS, (ii) regress p on x with OLS and obtain the residuals p^*, and (iii) construct the estimator as shown. The computational burden of the estimator is trivial, which may be especially beneficial in practice if it nested inside of a nonlinear routine for other parameters.[12]

We now develop a method-of-moments estimator that converges at the empirical root(s) of equation (11). Consider that the three-stage estimator rests on the moment condition $E[\xi' c] = 0$, which represents the combination of $E[\xi'\eta] = 0$, $E[X\xi] = 0$, and $E[\xi] = 0$. An alternative approach to estimation is to search numerically for a $\tilde{\beta}$ that satisfies the corresponding empirical moment, yielding

$$\hat{\beta}^{MM} = \arg\min_{\tilde{\beta}<0}\left[\frac{1}{T}\sum_t \frac{1}{|J_t|}\sum_{j\in J_t}\xi_{jt}(\tilde{\beta}; w, \sigma, X)\cdot c_{jt}(\tilde{\beta}; w, \sigma, X)\right]^2$$

where $\xi(\tilde{\beta}; w, \sigma, X)$ and $c(\tilde{\beta}; w, \sigma, X)$ are computed given the data and the candidate parameter using equations (5)-(7), and the firms present in each market t are indexed by the set J_t. The linear parameters (α, γ) are concentrated out of the nonlinear optimization problem. We have confirmed in numerical experiments that $\hat{\beta}^{\textit{3-Stage}}$ and $\hat{\beta}^{MM}$ are equivalent to numerical precision. Some care must be taken with the method-of-moments: if condition (12) fails then the optimizer reach the minimum (zero) at either the upper or lower root, and when the condition holds a local minimum may exist at the boundary value of the parameter space (as the optimizer attempts to reach the minimum for the positive root). Further, the three-stage estimator may immediately reject that a solution exists at the assumed value of $Cov(\xi, \eta)$, whereas the optimizer will return a solution.

There are three situations in which the method-of-moments approach may be preferred despite its greater computational burden. First, analytical solutions for $\frac{dh}{dq}$ may be unavailable with some specifications of the model, which diminishes the computation advantage of the three-stage estimator. Second, if valid instruments exist, then the additional moments suggest

[12] If condition (13) fails then the empirical analog to the upper root of equation (11) provides a consistent estimate. A more precise two-stage estimator is available for special cases in which the observed cost and demand shifters are uncorrelated. See Appendix C for details

a generalized method of moments estimator, allowing for efficiency improvements and specification tests (e.g., Hausman (1978); Hansen (1982)). Finally, the three-stage estimator requires orthogonality between the unobserved demand shock and *all* the regressors (i.e., $E[X\xi] = 0$). The method-of-moments approach can be pursued under a weaker assumption that allows for correlation between ξ and regressors that enter the cost function only. In this case, one would replace $E[\xi' c] = 0$ with $E[\xi' \eta] = 0$ in the objective function.

3.5 Small-Sample Properties

We generate Monte Carlo results to examine the small sample properties of the estimators. We consider a profit-maximizing monopolist that prices against a logit demand curve and has a constant marginal cost technology:

$$\begin{aligned} h(q_t; w_t) &\equiv \log(q_t) - \log(1 - q_t) = -\beta p_t + \xi_t \\ c_t &= x_t + \eta_t \end{aligned}$$

For simplicity, we set $\beta = 1$ and simulate data for x, ξ, η using independent $U[0, 1]$ distributions. For each draw of the data, we compute profit-maximizing prices and quantities. The mean price and margin are 2.20 and 0.56, respectively, and the mean price elasticity of demand is -1.86. We construct samples with 25, 50, 100, and 500 observations and estimate demand with each. We repeat this exercise 1,000 times and examine the average and standard deviation of the estimates. The estimators are the 3-Stage estimator, two-stage least squares (2SLS) using x_t as an instrument, a method-of-moments ("MM") estimator based on the alternative moment $E[\xi' \eta] = 0$, and OLS.

Table 2 summarizes the results. The bias present in 3-Stage, 2SLS, and MM is small even with the smallest sample sizes. However, 3-Stage more consistently provides accurate estimates than 2SLS and MM, as evidenced by the smaller standard deviation of the estimates. The reason is that 3-Stage utilizes orthogonality between unobserved demand and marginal cost, whereas 2SLS and MM exploit the relationship between unobserved demand and marginal cost shifters—either observed (x_t) or unobserved (η_t)—which provide noisy signals about marginal cost. One might be tempted to run a "first-stage" regression to test for the power of the different cost components to predict prices. However, such a test has no bearing on the asymptotic properties of the 3-Stage and MM estimators because exogenous supply-side variation need not be observed by the econometrician and indeed need not even exist. This is both a strength and a weakness: relaxing the requirement of observed exogenous variation comes at the cost of a greater reliance on assumptions about how firms set prices in equilibrium.

Table 2: Small Sample Properties of Estimators

Panel A: Average Estimates (Truth is $\beta = -1.00$)				
Sample Size	3-Stage	2SLS	MM	OLS
25	-1.002	-1.008	-1.005	-0.885
50	-1.004	-1.012	-1.002	-0.889
100	-1.004	-1.006	-1.005	-0.891
500	-1.000	-1.001	-0.999	-0.887
Panel B: Standard Deviation of Estimates				
Sample Size	3-Stage	2SLS	MM	OLS
25	0.160	0.276	0.208	0.168
50	0.109	0.182	0.141	0.114
100	0.078	0.123	0.101	0.082
500	0.035	0.053	0.045	0.037

Notes: The moments used for 3-Stage, 2SLS, MM, and OLS are $E[\xi' c]$, $E[\xi' x]$, $E[\xi' \eta]$, and $E[\xi' p]$, respectively. The methods-of-moments ("MM") estimator is implemented with a one-dimensional grid search.

4 Generalizations

The results developed thus far rely on an accurate model of the data generating process and some relatively strong (though common) restrictions on the form of demand and supply. In this section, we consider generalizations to non-constant marginal costs, multi-product firms, and non-Bertrand competition.

4.1 Non-Constant Marginal Costs

If marginal costs are not constant in output, then unobserved demand shocks that change quantity also affect marginal cost. For example, consider a special case in which marginal costs take the form:

$$c_{jt} = x'_{jt}\gamma + g(q_{jt}; \lambda) + \eta_{jt} \tag{14}$$

Here $g(q_{jt}; \lambda)$ is some potentially nonlinear function that may (or may not) be known to the econometrician. Maintaining Bertrand competition and the baseline demand assumption, the first-order conditions of the firm are:

$$p_{jt} = \underbrace{x'_{jt}\gamma + g(q_{jt}; \lambda) + \eta_{jt}}_{\text{Marginal Cost}} + \underbrace{\left(-\frac{1}{\beta}\frac{dh_{jt}}{dq_{jt}} q_{jt}\right)}_{\text{Markup}}.$$

Thus, provided $g'(\cdot; \lambda) \neq 0$, markup adjustments are no longer the only mechanism through with prices respond to demand shocks. Unless knowledge of $g(q_{jt}; \lambda)$ can be brought to bear, the identification results of the preceding section do not extend without additional restrictions.

This also can be seen from the OLS regression of $h(q_{jt}, w_{jt}; \sigma)$ on p and x, which yields a price coefficient with the following probability limit:

$$plim(\hat{\beta}^{OLS}) = \beta - \frac{1}{\beta}\frac{Cov(\xi, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov(\xi, g(q))}{Var(p^*)}$$

The third term on the right-hand-side shows that bias depends on how demand shocks affect the non-constant potion of marginal costs.

There are two ways to make progress. First, if $g'(\cdot;\lambda)$ can be signed then it is possible to bound the price parameter, β, even if point identification remains infeasible. A lead example is that of capacity constraints, for which it might be reasonable to assume that $Cov(\xi, \eta) = 0$ and $g'(\cdot;\lambda) \geq 0$, and thus that $Cov(\xi, \eta^*) \geq 0$ where $\eta^*_{jt} \equiv \eta_{jt} + g(q_{jt};\lambda)$ is a composite error term. Bounds with priors then can be constructed. Second, the econometrician may be able to estimate $g(q_{jt};\lambda)$, either in advance or simultaneously with the price coefficient. Prior knowledge of $Cov(\xi, \eta)$ is sufficient to at least set identify β in such a situation:

Proposition 6. *Under assumptions 1 and 3 and a modified assumption 2 in which marginal costs take the semi-linear form of equation* (14)*, the price parameter β solves the following quadratic equation:*

$$\begin{aligned} 0 &= \left(1 - \frac{Cov(p^*, g(q))}{Var(p^*)}\right)\beta^2 \\ &+ \left(\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov(\xi, \eta)}{Var(p^*)} - \hat{\beta}^{OLS} + \frac{Cov(p^*, g(q))}{Var(p^*)}\hat{\beta}^{OLS} + \frac{Cov(\hat{\xi}^{OLS}, g(q))}{Var(p^*)}\right)\beta \\ &+ \left(-\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\hat{\beta}^{OLS} - \frac{Cov(\hat{\xi}^{OLS}, \frac{dh}{dq}q)}{Var(p^*)}\right) \end{aligned}$$

where $\hat{\beta}^{OLS}$ is the OLS estimate and $\hat{\xi}^{OLS}$ is a vector containing the OLS residuals.

Proof. See appendix.

With the above quadratic in hand, the remaining results of Section 3 extend naturally. Although the estimation of $g(q_{jt};\lambda)$ is not our focus, we note that a three-stage estimator of β could be obtained for any candidate parameters in λ, thereby facilitating computational efficiency.

4.2 Multi-Product Firms

We now provide the notation necessary to extend our results to the case of multi-product firms under our maintained assumptions. Let K^m denote the set of products owned by multi-product firm m. When the firm sets prices on each of its products to maximize joint profits, there are

$|K^m|$ first-order conditions, which can be expressed as

$$\sum_{k \in K^m} (p_k - c_k) \frac{\partial q_k}{\partial p_j} = -q_j \;\; \forall j \in K^m.$$

The market subscript, t, is omitted to simplify notation. For demand systems satisfying Assumption 1,

$$\frac{\partial q_k}{\partial p_j} = \beta \frac{1}{\frac{dh_j}{dq_k}}.$$

where the derivative $\frac{dh_j}{dq_k} = \frac{\partial h_j}{\partial q_j}\frac{dq_j}{dq_k} + \frac{\partial h_j}{\partial w_j}\frac{dw_j}{dq_k}$ is calculated holding the prices of other products fixed. Therefore, the set of first-order conditions can be written as

$$\sum_{k \in K^m} (p_k - c_k) \frac{1}{dh_j/dq_k} = -\frac{1}{\beta} q_j \;\; \forall j \in K^m.$$

Stack the first-order conditions, writing the left-hand side as the product of a vector of markups $(p_j - c_j)$ and a matrix A^m of loading components, $A^m_{i(j),i(k)} = \frac{1}{dh_j/dq_k}$, where $i(\cdot)$ indexes products within a firm. Next, invert the loading matrix to solve for markups as function of the loading components and $-\frac{1}{\beta}\boldsymbol{q}^m$, where $\boldsymbol{q}^m$ is a vector of the multi-product firm's quantities. Equilibrium prices equal marginal costs plus a markup, where the markup is determined by the inverse of A^m $((A^m)^{-1} \equiv \Lambda^m)$, quantities, and the price parameter:

$$p_j = c_j - \frac{1}{\beta} \left(\Lambda^m \boldsymbol{q}^m\right)_{i(j)}. \tag{15}$$

Here, $\left(\Lambda^m \boldsymbol{q}^m\right)_{i(j)}$ provides the entry corresponding to product j in the vector $\Lambda^m \boldsymbol{q}^m$. As the matrix Λ^m is not a function of the price parameter after conditioning on observables, this form of the first-order condition allows us to solve for β using a quadratic three-stage solution analogous to that in equation (2).[13] Letting $\tilde{h} \equiv \left(\Lambda^m \boldsymbol{q}^m\right)_{i(j)}$ be the multi-product analog for $\frac{dh}{dq}q$, we obtain a quadratic in β, and the remaining results of Section 3 then obtain easily:

Corollary 3. *Under assumptions 1 and 3, along with a modified assumption 2 that allows for multi-product firms, the price parameter β solves the following quadratic equation:*

$$\begin{aligned} 0 &= \beta^2 \\ &+ \left(\frac{Cov(p^*, \tilde{h})}{Var(p^*)} + \frac{Cov(\xi, \eta)}{Var(p^*)} - \hat{\beta}^{OLS}\right)\beta \\ &+ \left(-\frac{Cov(p^*, \tilde{h})}{Var(p^*)}\hat{\beta}^{OLS} - \frac{Cov(\hat{\xi}^{OLS}, \tilde{h})}{Var(p^*)}\right). \end{aligned}$$

[13] At this point, the reader may be wondering where the prices of other firms are captured under the adjusted first-order conditions for multi-product ownership. As is the case with single product firms, we expect prices of other firm's products to be included in w_j, which is appropriate under Bertrand price competition.

where $\tilde{h}$ is constructed from the first-order conditions of multi-product firms.

4.3 Alternative Models of Competition

Though our main results are presented under Bertrand competition in prices, our method applies to a broader set of competitive assumptions. Consider, for example, Nash competition among profit-maximizing firms that have a single choice variable, a, and constant marginal costs. The individual firm's objective function is:

$$\max_{a_j | a_i, i \neq j} (p_j(a) - c_j) q_j(a).$$

This generalized model of Nash competition nests Bertrand ($a = p$) and Cournot ($a = q$). The first-order condition, holding fixed the actions of the other firms, is given by:

$$p_j(a) = c_j - \frac{p_j{}'(a)}{q_j{}'(a)} q_j(a).$$

In equilibrium, we obtain the structural decomposition $p = c + \mu$, where μ incorporates the structure of demand and its parameters. This decomposition provides a restriction on how prices move with demand shocks, aiding identification. It can be obtained in other contexts, including consistent conjectures and competition in quantities with increasing marginal costs.[14] When the markup is proportional to the reciprocal of the price parameter, then it is straightforward to extend our core identification result and implement the three-stage estimator. We provide one such extension in the empirical application to the cement industry.

5 Empirical Applications

5.1 The Portland Cement Industry

Our first empirical application uses the setting and data of Fowlie et al. (2016) ["FRR"], which examines market power in the cement industry and its effects on the efficacy of environmental regulation. The model features Cournot competition among undifferentiated cement plants facing capacity constraints.[15] As we describe below, institutional details about cement demand and the production process support the reasonableness of uncorrelatedness in the model.

We begin by extending our results to Cournot competition with non-constant marginal costs. Let $j = 1, \ldots, J$ firms produce a homogeneous product demanded by consumers according to

[14]Nonetheless, some models are excluded. For example, a monopolist facing a log-linear demand schedule sets prices according to $p = c\frac{\epsilon}{1+\epsilon}$ where $\epsilon < 0$ is the elasticity of demand.

[15]A published report of the Environment Protection Agency (EPA) states that "consumers are likely to view cement produced by different firms as very good substitutes.... there is little or no brand loyalty that allows firms to differentiate their product" EPA (2009).

$h(Q;w) = \beta p + x'\gamma + \xi$, where $Q = \sum_j q_j$, and p represents a price common to all firms in the market. Marginal costs are semi-linear, as in equation (14), possibly reflecting capacity constraints. Working with aggregated first-order conditions, it is possible to show that the OLS regression of $h(Q;w_{jt})$ on price and covariates yields:

$$plim(\hat{\beta}^{OLS}) = \beta - \frac{1}{\beta}\frac{1}{J}\frac{Cov(\xi, \frac{dh}{dq}Q)}{Var(p^*)} + \frac{Cov(\xi, \overline{g})}{Var(p^*)}$$

where J is the number of firms in the market and $\overline{g} = \frac{1}{J}\sum_{j=1}^{J} g(q_j;\lambda)$ is the average contribution of $g(q,\lambda)$ to marginal costs. Bias arises due to markup adjustments and the correlation between unobserved demand and marginal costs generated through $g(q;\lambda)$.[16] The identification result provided in Section 4.1 for models with non-constant marginal costs extends.

Corollary 4. *In the Cournot model, the price parameter β solves the following quadratic equation:*

$$\begin{aligned}
0 &= \left(1 - \frac{Cov(p^*, \overline{g})}{Var(p^*)}\right)\beta^2 \\
&+ \left(\frac{1}{J}\frac{Cov(p^*, \frac{dh}{dq}Q)}{Var(p^*)} + \frac{Cov(\xi, \overline{\eta})}{Var(p^*)} - \hat{\beta}^{OLS} + \frac{Cov(p^*, \overline{g})}{Var(p^*)}\hat{\beta}^{OLS} + \frac{Cov(\hat{\xi}^{OLS}, \overline{g})}{Var(p^*)}\right)\beta \\
&+ \left(-\frac{1}{J}\frac{Cov(p^*, \frac{dh}{dq}Q)}{Var(p^*)}\hat{\beta}^{OLS} - \frac{1}{J}\frac{Cov(\hat{\xi}^{OLS}, \frac{dh}{dq}Q)}{Var(p^*)}\right)
\end{aligned}$$

The derivation tracks exactly the proof of Proposition 6. For the purposes of the empirical exercise, we compute the three-stage estimator as the empirical analog to the lower root.

Turning to the application, FRR examine 20 distinct geographic regions in the United States annually over 1984-2009. Let the demand curve in region r and year t have a logit form:

$$h(Q_{rt};w) \equiv \ln(Q_{rt}) - \ln(M_r - Q_{rt}) = \alpha_r + \beta p_{rt} + \xi_{rt}$$

where M_r is the "market size" of the region. We assume $M_r = 2 \times \max_t\{Q_{rt}\}$ for simplicity.[17] Further, let marginal costs take the "hockey stick" form of FRR:

$$\begin{aligned}
c_{jrt} &= \gamma + g(q_{jrt}) + \eta_{jrt} \\
g(q_{jrt}) &= 2\lambda_2 1\{q_{jrt}/k_{jr} > \lambda_1\}(q_{jrt}/k_{jr} - \lambda_1)
\end{aligned}$$

where k_{jr} and q_{jrt}/k_{jr} are capacity and utilization, respectively. Marginal costs are constant

[16] Bias due to markup adjustments dissipates as the number of firms grows large. Thus, if marginal costs are constant then the OLS estimate is likely to be close to the population parameter in competitive markets. In Monte Carlo experiments, we have found similar results for Bertrand competition and logit demand.

[17] We use logit demand rather than the constant elasticity demand of FRR because it fits easily into our framework. The 2SLS results are unaffected by the choice. Similarly, the 3-Stage estimator with logit obtains virtually identical results as a method-of-moments estimator with constant elasticity demand that imposes uncorrelatedness.

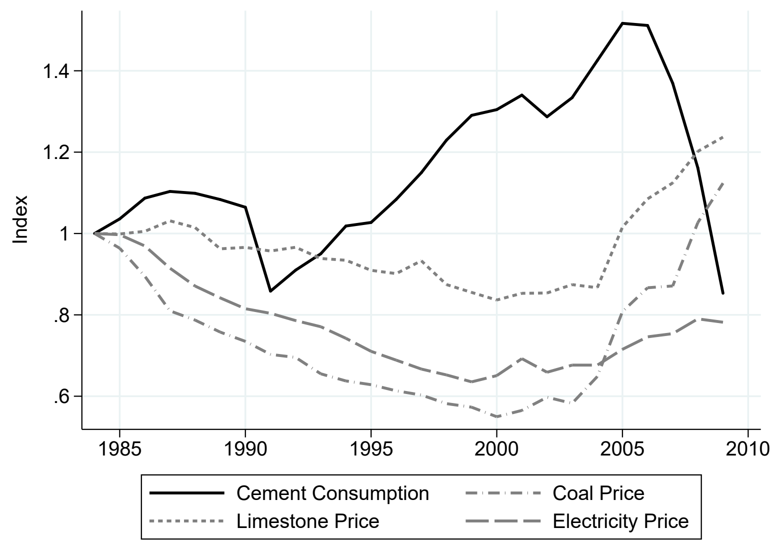

Figure 2: Statistics on the Cement Industry

Notes: Total cement consumption in the United States and national average limestone prices are from the *Minerals Yearbook* of the United States Geological Survey. National average coal prices for industrial users are from the State Energy Database System of the Energy Information Administration (EIA). National average electricity prices for industrial users are from from the Annual Energy Review of the EIA. All data are annual except the limestone price, which is available in even years through the 1980s and annually thereafter. Prices are adjusted for inflation with the GDP Deflator.

if utilization is less than the threshold $\lambda_1 \in [0, 1]$, and increasing linearly at rate determined by $\lambda_2 \geq 0$ otherwise. The two unobservables, (ξ, η), capture demand shifts and shifts in the constant portion of marginal costs.

The institutional details of the industry suggest that uncorrelatedness may be reasonable. Demand is procyclical because cement is used in construction projects; given the demand specification this cyclicality enters through the unobserved demand shock. On the supply side, the two largest cost components are "materials, parts, and packaging" and "fuels and electricity" (EPA, 2009). Both depend on the price of coal. With regard to "fuels and electricity," most cement plants during the sample period rely on coal as their primary fuel, and electricity prices are known to correlate with coal prices. With regard to "material, parts, and packaging," the main input in cement manufacture is limestone, which requires significant amounts of electricity to extract (National Stone Council, 2008). Thus, an assessment of uncorrelatedness hinges largely on the relationship between construction activity and coal prices.

In this context, there is a theoretical basis for orthogonality: if coal suppliers have limited market power and roughly constant (realized) marginal costs, then coal prices should not respond much to demand. To explore this possibility, Figure 2 plots cement consumption, limestone prices, coal prices, and electricity prices. Consumption, which has a tight empirical

Table 3: Point Estimates for Cement

Estimator:	3-Stage	2SLS	OLS
Elasticity	-1.15	-1.07	-0.47
of Demand	(0.18)	(0.19)	(0.14)

Notes: The sample includes 520 region-year observations over 1984-2009. Bootstrapped standard errors are based on 200 random samples constructed by drawing regions with replacement.

connection to contruction activity, exhibits the aforementioned procyclicality.[18] The cost statistics, by contrast, decrease gradually over 1984-2003 and then increase over 2004-2009. The stark differences between the consumption and input price patterns support an assumption along the lines of $Cov(\xi, \overline{\eta}) = 0$. Indeed, this is precisely the identification argument of FRR, as both coal and electricity prices are included in the set of excluded instruments.[19]

Table 3 summarizes the results of demand estimation. The 3-Stage estimator is implemented taking as given the nonlinear cost parameters obtained in FRR: $\lambda_1 = 0.869$ and $\lambda_2 = 803.65$. In principle, these could be estimated simultaneously via the method of moments, provided some demand shifters can be excluded from marginal costs, but estimation of these parameters is not our focus. As shown, the mean price elasticity of demand obtained with the 3-Stage estimator under uncorrelatedness is -1.15. This is statistically indistinguishable from the 2SLS elasticity estimate of -1.07, which is obtained using the FRR instruments: coal prices, natural gas prices, electricity prices, and wage rates. The closeness of the 3-Stage and 2SLS is not coincidental and instead reflects that the identifying assumptions are quite similar. Indeed, the main difference is whether the cost shifters are treated as observed (FRR) or unobserved (3-Stage).

If the econometrician does not know (and cannot identify) the nonlinear parameters in the cost function, then consistent estimates cannot be obtained with our methodology. Further, one can confirm that prior-free bounds are not available as the empirical upper root of the quadratic in Corollary 4 is positive. Nonetheless, some progress can be made using posterior bounds. Define the composite marginal cost shock, $\eta^*_{jrt} = g(q_{jrt}) + \eta_{jrt}$, as inclusive of the capacity effects. Given the upward-sloping marginal costs, we have $Cov(\xi, \overline{\eta}^*) \geq 0$ if $Cov(\xi, \overline{\eta}) = 0$. This restriction generates an upper bound on the demand elasticity of -0.69, ruling out the OLS point estimate.

[18]Macher et al. (2018) report that data on construction employment and building permits are sufficient to explain 90 percent of the variation in state-level consumption.

[19]A close examination of Figure 2 may suggest a *negative* empirical correlation between the demand and cost measures, and indeed this is the case: the correlation coefficient between consumption and coal prices is -0.20, for example. One interpretation consistent with uncorrelatedness is that this arises due to the short sample. After materials and fuel, the third largest cost component is labor, which may be partially fixed (EPA, 2009). The decline of union power during the Reagan administration led to declining labor costs through the 1980s (Dunne et al., 2009). FRR employ wage rates as an instrument.

5.2 The Airline Industry

In our second empirical exercise, we estimate demand for airline travel using the setting and data of Aguirregabiria and Ho (2012) ["AH"].[20] AH explores why airlines form hub-and-spoke networks; here, we focus on demand estimation only. The model features differentiated-products Bertrand competition among multi-product firms facing a nested logit demand system. To identify the demand system, we employ the multi-product variant of our theoretical results developed in Section 4.2, proceeding under the assumption that β is obtained at the lower root. We provide point estimates for demand parameters under uncorrelatedness and demonstrate how weaker assumptions can be used to set identify key parameters.

The nested logit demand system can be expressed as

$$h(s_{jmt}, w_{jmt}; \sigma) \equiv \ln s_{jmt} - \ln s_{0mt} - \sigma \ln \overline{s}_{jmt|g} = \beta p_{jmt} + x'_{jmt}\alpha + \xi_{jmt} \tag{16}$$

where s_{jmt} is the market share of product j in market m in period t. The conditional market share, $\overline{s}_{j|g} = s_j / \sum_{k \in g} s_k$, is the the choice probability of product j given that its "group" of products, g, is selected. The outside good is indexed as $j = 0$. Higher values of σ increase within-group consumer substitution relative to across-group substitution. In contrast to the typical expression for the demand system, we place $\sigma \ln \overline{s}_{jmt|g}$ on the left-hand side so that the right-hand side contains a single endogenous regressor: price.

Equation (16) results from a standard discrete-choice utility formulation where consumers have correlated preferences for products within the same group. In the airline setting, markets are directional round trips between origin and destination cities in a particular quarter. Consumers within a market choose among airlines and whether to take a nonstop or one-stop itinerary. Thus, each airline offers zero, one, or two products per market. The nesting parameter, σ, governs consumer substitution within each product group: nonstop flights, one-stop flights, and the outside good. Marginal costs are linear in accordance with equation (6).

The data are drawn from the *Airline Origin and Destination Survey* (DB1B) survey, a ten percent sample of airline itineraries, for the four quarters of 2004. Following AH, the covariates include an indicator for nonstop itineraries, the distance between the origin and destination cities, and a measure of the airline's "hub sizes" at the origin and destination cities. We also include airline fixed effects and route$\times$quarter fixed effects. The latter expands on the city$\times$quarter fixed effects described by AH. Market size, which determines the market share of the outside good, is equal to the total population in the origin and destination cities.

[20]We thank Victor Aguirregabiria for providing the data. Replication is not exact because the sample differs somewhat from what is used in the AH publication and because we employ a different set of fixed effects in estimation.

Point Identification and Estimation

We now consider identification of the nesting parameter, σ. In the nested logit model, conditional shares respond to the unobserved demand shock, which creates a second endogenous variable in addition to price. As the assumptions generating the three-stage estimator provide a single restriction, additional moments are required to pin down β and σ simultaneously. Uncorrelatedness alone provides an identified set through a function that maps σ to β. For the purpose of discussing point identification, we take uncorrelatedness as given, and we return to the subject when we discuss bounds.

Combining a single supplemental moment, $E[f] = 0$, with the uncorrelatedness assumption $E[\xi'\eta] = 0$ is sufficient to identify the additional parameter.[21] Provided these two moments, the econometrician could pursue a method-of-moments estimator, searching over the parameter space for the pair (β,σ) that minimizes a weighted sum of the moments (squared). Since the three-stage estimator is consistent for β conditional on σ, the econometrician could instead conduct a single-dimensional search, obtaining $\hat{\beta}^{3-Stage}(\tilde{\sigma})$ for each candidate parameter $\tilde{\sigma}$ while minimizing the objective $(E[f] - 0)^2$. As this approach has a distinct computational advantage, we follow it throughout this section.

The supplemental moments $E[f] = 0$ may be a vector and can be obtained from standard instrumental variable assumptions ($E[Z\xi] = 0$). To construct these moments, we use covariates from the AH marginal cost function that are excluded from demand.[22] These instruments differ from the AH instruments included in 2SLS estimation, which we describe below.

In addition to supplementing the covariance restriction with instruments, the econometrician could derive additional covariance restrictions based on reasonable extensions of the notion of uncorrelatedness. For example, if product-level shocks are uncorrelated, it may be reasonable to assume that shocks are uncorrelated when aggregated by product group. This approach does not require the econometrician to be able to isolate exogenous variation in prices and conditional shares. We consider two supplementary covariance restrictions:

- $Cov(\overline{\xi}, \overline{\eta}) = 0$ where $\overline{\xi}_{gt} = \frac{1}{|g|}\sum_{j\in g}\xi_{jt}$ and $\overline{\eta}_{gt} = \frac{1}{|g|}\sum_{j\in g}\eta_{jt}$ are the mean demand and cost shocks within a group-market pair. This is a simple refinement of uncorrelatedness: the mean shocks within a product group are uncorrelated across groups and markets.

- $Cov(\xi^2, \eta) = 0$ and $Cov(\xi, \eta^2) = 0$. These identifying assumptions state that the variance of one shock is uncorrelated with the level of the other shock.

Note that the latter assumption does not provide independent identifying power if (ξ, η) are jointly normal, because it would be implied by orthogonality.

[21] We assume the moment meets the necessary support conditions.

[22] In the demand equation, hub size of any given city-airline pair is the sum of population in other cities that the airline connects with direct itineraries from the city. In the supply equation, this is replaced with an analogous measure based on the number of connections rather than population.

Table 4: Application to U.S. Airlines

Parameter	3-Stage-I	3-Stage-II	3-Stage-III	2SLS	OLS
β	-0.182	-0.153	-0.131	-0.189	-0.106
	(0.042)	(0.031)	(0.031)	(0.053)	(0.004)
σ	0.525	0.599	0.639	0.822	0.891
	(0.110)	(0.113)	(0.140)	(0.087)	(0.003)

Notes: The first two columns of results use the three-stage methodology with different supplemental moments. The next column use 2SLS and OLS, respectively, following Aguirregabiria and Ho (2012). Standard errors are constructed via subsamples of 100 market-periods. There are 93,199 observations and 11,474 market-periods in the full sample.

Table 4 summarizes the results of estimation. The left three columns are obtained with the three-stage methodology sketched above, using first $E[Z\xi] = 0$ as a supplemental moment, then using $Cov(\overline{\xi}, \overline{\eta}) = 0$, and finally using $Cov(\xi^2, \eta) = 0$ and $Cov(\xi, \eta^2) = 0$. The fourth column is obtained with 2SLS using the AH instruments that are not absorbed by route$\times$quarter fixed effects: the average hub-sizes (origin and destination) of all other airlines on the route and the average value of the nonstop indicator for all the other carriers on the route. The final column is obtained with OLS. The 3-Stage estimators and 2SLS all move the parameters in the expected direction relative to OLS. Comparing the estimates, 3-Stage produces less negative price parameters and smaller nesting parameters than 2SLS. We do not seek to ascertain which set of estimates is more in line with real-world behavior.[23]

Bounds and Set Identification

In many settings, the econometrician might prefer to proceed under weaker assumptions about the correlation between unobservables in a structural model. As discussed earlier, our results can be used to place bounds on the parameter space. Without imposing additional restrictions on the correlation structure, the econometrician can obtain prior-free bounds to reject parameter values that are inconsistent with the observed data, conditional on the model.[24] To further narrow the identified set, the econometrician can invoke knowledge of institutional details of the industry under study.

For an airline, the marginal cost of an additional passenger is small and roughly constant until the plane nears capacity. Each additional passenger has a little impact on the inputs needed to fly the plane from one airport to another. However, the airline bears an opportunity

[23]Ciliberto et al. (2016) partially identify a correlation coefficient of $Cor(\xi, \eta) \in [0.38, 0.40]$ based on similar data from 2012, and this potentially calls into question the reasonableness of the uncorrelatedness assumption in the airlines industry. Alternatively, their result could be an artifact of the demand specification, which does not incorporate fixed effects.

[24]For an illustration of the link between rejected values of $Cov(\xi, \eta)$ and rejected values of β in this application, see Figure E.1 in the Appendix.

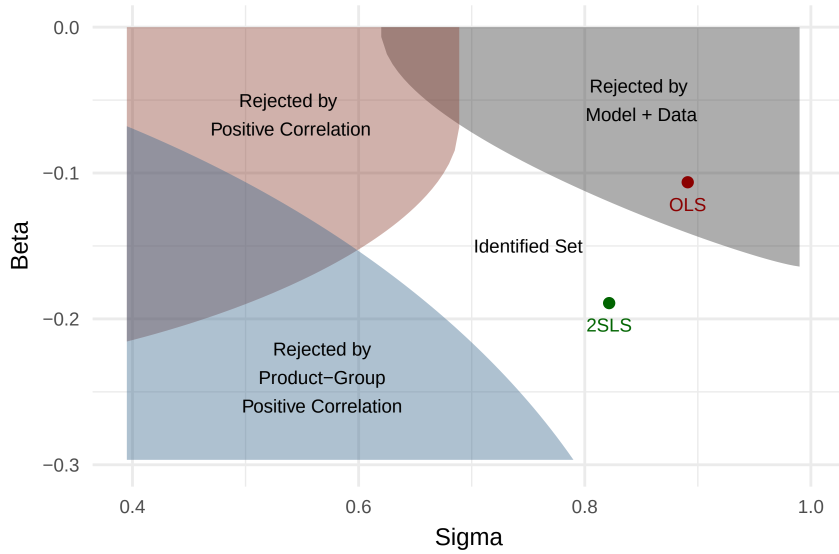

Figure 3: Identified Parameter Set Under Priors

Notes: Figure displays candidate parameter values for (σ, β). The gray region indicates the set of parameters that cannot generate the observed data from the assumptions of the model. The red region indicates the set of parameters that generate $Cov(\xi, \eta) < 0$, and the blue region indicates parameters that generate $Cov(\bar{\xi}, \bar{\eta}) < 0$. The identified set is obtained by rejecting values in the above regions under the assumption of (weakly) positive correlation. For context, the OLS and the 2SLS estimates are plotted, along with 3-Stage estimates I-IV. The parameter σ can only take values on $[0, 1)$.

cost for each sold seat, as they can no longer sell the seat at a higher price to another passenger (Williams, 2017). When only a few seats are left to sell, this opportunity cost may become large, approximating the "hockey-stick" cost function that describes the cement industry. As positive shocks to demand result in more full flights, it is reasonable to assume that the correlation between demand shocks and marginal cost shocks is weakly positive across markets and also for individual firms.[25]

Under this assumption, it is possible to reject values (β,σ) that produce negative correlation. The econometrician can combine this prior with the prior-free bounds developed earlier. Finally, one can consider reasonable extensions of the priors over the correlation between demand and supply shocks. Following the logic above, if the correlation in product-level shocks is weakly positive, one can assume that the group-level shocks are also weakly positive. By rejecting parameter values that fail to generate the data or that deliver negative correlation, the econometrician can narrow the identified set.

Figure 3 displays the rejected regions based on the model and above priors on unobserved shocks. The gray region corresponds to the parameter values rejected by the prior-free bounds.

[25] The shocks many not be strictly positive, given the fixed effects we included in demand and the fact that the data are aggregated to the quarterly level.

Some values of β can be rejected if $\sigma \geq 0.62$. As σ becomes larger, a more negative β is required to rationalize the data within the context of the model. If $\sigma = 0.80$ then it must be that $\beta \leq -0.11$. The dark red region corresponds to parameter values that generate negative correlation between demand and supply shocks. These can be rejected under the prior that $Cov(\xi, \eta) \geq 0$. The dark blue region provides the corresponding set for the prior $Cov(\overline{\xi}, \overline{\eta}) \geq 0$.

The three regions overlap, but no region is a strict subset of the other. In this manner, the econometrician can impose additional restrictions on the covariance structure to rule out parameter values. The remaining non-rejected values provide the identified set. In our application, we are able to rule out values of σ less than 0.599 for any value of β, as these lower values cannot generate positive correlation in both product-level and product-group-level shocks. Similarly, we obtain an upper bound on β of -0.067 across all values of σ.

For context, we plot the OLS and the 2SLS estimates in Figure 3. The OLS estimate falls in a rejected region and can be ruled out by the structure of the model alone. The 2SLS estimate, in contrast, falls within the identified set. As the assumptions underlying the bounds do not correspond to the IV assumptions, it is possible that IV point estimates may be ruled out with this approach in other applications.

6 Discussion

6.1 Assessing Covariance Restrictions

In many applications, econometricians may have detailed knowledge of the determinants of demand and marginal cost, even if many determinants are unobserved in the data. Such knowledge allows the econometrician to assess whether covariance restrictions along the lines of $Cov(\xi, \eta) = 0$ are reasonable. Covariance restrictions need not (and should not) be a "black box" that provides identification. The distinction between *observed* and *understood* is important, as the econometrician may have reasonable priors about the relationships between structural error terms even though they are (by definition) unobserved. For a constructive example, see the discussion about the cement industry in Section 5.1.

Sometimes knowledge of institutional details suggests that uncorrelatedness may be *unreasonable* as an identifying assumption. Products with greater unobserved quality might be more expensive to produce, demand shocks could raise or lower marginal costs (e.g., due to capacity constraints), or firms might invest to lower the costs of their best-selling products. In these cases, 3-Stage estimates under uncorrelatedness would be inconsistent unless the confounding variation can be absorbed by control variables or fixed effects. Rich panel data provides the econometrician with the means to correct for several first-order determinants of correlation, as we show below. Even without these controls, it may be possible to sign $Cov(\xi, \eta)$, allowing the econometrician to set identify parameters using bounds with priors. As with any identification strategy, careful attention must be paid to the institutional details.

We highlight that econometricians with panel data may be able to incorporate fixed effects that absorb otherwise confounding correlations. To illustrate, consider the following generalized demand and cost functions:

$$\begin{aligned} h(q_{jt}, w_{jt}; \sigma) &= \beta p_{jt} + x'_{jt}\alpha + D_j + F_t + E_{jt} \\ c_{jt} &= g(q_{jt}; \lambda) + x'_{jt}\gamma + U_j + V_t + W_{jt} \end{aligned}$$

with $Cov(E_{jt}, W_{jt}) = 0$. Let the unobserved shocks be $\xi_{jt} = D_j + F_t + E_{jt}$ and $\eta_{jt} = U_j + V_t + W_{jt}$. Further, let $h(\cdot)$ and $g(\cdot)$ be known up to parameters. If products with higher quality have higher marginal costs then $Cov(U_j, D_j) > 0$. The econometrician can account for the relationship by estimating D_j for each firm; the residual $\xi^*_{jt} = \xi_{jt} - D_j$ is uncorrelated with U_j. Similarly, if costs are higher (or lower) in markets with high demand then market fixed effects can be incorporated. In this manner, panel data reduce the remaining unobserved correlation to product-specific deviations within a market, E_{jt} and W_{jt}, allowing the econometrician to proceed with the three-stage approach. Of course, the econometrician must assess whether the restriction $Cov(E_{jt}, W_{jt}) = 0$ is appropriate in the empirical setting.

6.2 Relation to Instruments

To further build intuition on covariance restrictions, we draw some connections between estimation under uncorrelatedness and the instrumental variation approach. In our view, the most obvious similarity is that both approaches rely on orthogonality conditions—$E[\xi'\eta]$ or $E[\xi'Z]$ for instruments Z—that are not verifiable empirically but can be assessed with knowledge of institutional details. A stylized model makes this connection clear: Suppose marginal costs are determined by a single variable w that is orthogonal to the demand-side structural error term. If the variable is observed, then IV estimation can proceed under $E[\xi'Z]$ with $Z = w$ and if it is unobserved than $E[\xi'\eta]$ with $\eta = w$ allows for estimation via uncorrelatedness. Indeed, some existing articles on covariance restrictions refer to the supply-side structural error term as providing an "unobserved instrument" that identifies demand (e.g., Hausman and Taylor (1983); Matzkin (2016)).[26]

In general, the assumptions embedded by the two approaches are not nested. Consider the case where marginal costs are the sum of an observed and unobserved variable: $c = Z + \eta$. When $Cov(\eta, \xi) \neq 0$, the IV conditions may still be satisfied, whereas three-stage estimation requires both Z and η to be orthogonal to ξ. On the other hand, the conditions needed for consistent three-stage estimation are not sufficient for IV, as IV requires that Z *does not enter the*

[26] Hausman and Taylor (1983) proposes a two-stage approach for the estimation of supply and demand models of perfect competition: First, the supply equation is estimated with 2SLS using an instrument taken from the demand-side of the model. Second, the supply-side error term is recovered and, under the assumption of uncorrelatedness ($Cov(\xi, \eta) = 0$), it serves as a valid instrument for the estimation of demand.

demand equation.[27] Satisfying this assumption is one of the key challenges in finding a plausible instrument and is not necessary for estimation under uncorrelatedness. In addition to this theoretical distinction, the IV approach has an additional empirical requirement related to the weak instruments problem: the instrumental variables approach requires sufficient variation in the observed instrument Z, whereas estimation under uncorrelatedness can proceed even if the cost determinants (Z and η) exhibit no variation.

Finally, estimation with uncorrelatedness requires a correctly-specified supply side, whereas IV requires a less formal theory of supply. Of course, many research articles that estimate demand maintain supply-side specifications for counterfactual experiments, and some articles also use supply-side moments to improve efficiency in demand estimation (e.g., Berry et al. (1995)). Nonetheless, researchers sometimes express a preference for demand to estimated separately, which ensures that at least the demand estimates are not influenced by misspecification on the supply side (e.g., Dubé and Chintagunta (2003)). Econometricians relying on uncorrelatedness for identification do not have that option—demand and supply must be estimated jointly, increasing the importance of efforts to validate the supply-side assumptions.

7 Conclusion

Our objective has been to evaluate the identifying power of typical supply-side restrictions in models of imperfect competition. Our main result is that price endogeneity can be resolved by interpreting an OLS estimate through the lens of a theoretical model. With a covariance restriction, the demand system is point identified, and weaker assumptions generate bounds on the structural parameters. Thus, causal demand parameters can be recovered without the availability of exogenous price variation. We hope that the methods we introduce help facilitate research in areas for which strong instruments are unavailable or difficult to find.

Though we focus our results on specific, widely-used assumptions about demand and supply, we view our method as not particular to these assumptions. Rather, the main insight is that information about supply-side behavior can be modeled to adjust the observed relationships between quantity and price. Price can be decomposed into marginal cost and a markup; our method provides a direct way to correct for endogeneity arising from the latter component. In a more general sense, this insight has a similar flavor to control function estimation procedures (e.g., Heckman (1979)). Our method may be thought of a bias correction procedure for empirical applications with models of imperfect competition.

[27]If it does, it would violate either the relevance condition or the exclusion restriction, depending on interpretation.

References

Aguirregabiria, V. and C.-Y. Ho (2012, May). A dynamic oligopoly game of the US airline industry: Estimation and policy experiments. *Journal of Econometrics 168*(1), 156–173.

Bajari, P. (2003). Comment on "Bayesian analysis of simultaneous demand and supply". *Quantitative Marketing and Economics 1*, 277–283.

Berry, S. (2003). Comment on "Bayesian analysis of simultaneous demand and supply". *Quantitative Marketing and Economics 1*, 285–291.

Berry, S. and P. A. Haile (2014). Identification in differentiated products markets using market level data. *Econometrica 82*(5), 1749–1797.

Berry, S., J. Levinsohn, and A. Pakes (1995, July). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics 25*(2), 242–262.

Bresnahan, T. (1987). Competition and collusion in the american automobile oligopoly: The 1955 price war. *Journal of Industrial Economics 35*(4), 457–482.

Bresnahan, T. (1996). Valuation of new goods under perfect and imperfect competition: Comment. In *The Economics of New Goods, Studies in Income and Wealth* Vol 58, ed. by T. Bresnahan and R. Gordon. Chicago: National Bureau of Economic Research.

Broda, C. and D. E. Weinstein (2006). Globalization and the gains from variety. *The Quarterly journal of economics 121*(2), 541–585.

Byrne, D. P., S. Imai, N. Jain, V. Sarafidis, and M. Hirukawa (2016). Identification and estimation of differentiated products models using market size and cost data.

Cardell, S. N. (1997). Variance components structures for the extreme-value and logistic distributions with applications to models of heterogeneity. *Econometric Theory 13*, 185–213.

Chiappori, P.-A., I. Komunjer, and D. Kristensen (2017). Nonparametric identification in nonlinear simultaneous equations models: The case of covariance restrictions. *Working Paper*.

Cho, S., G. Lee, J. Rust, and M. Yu (2018). Optimal dynamic hotel pricing.

Ciliberto, F., C. Murry, and E. Tamer (2016). Market structure and competition in airline markets.

Crawford, G. S. and A. Yurukoglu (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review 102*(2), 643–685.

De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica 79*(3), 1407–1451.

Deaton, A. and J. Muellbauer (1980). An almost ideal demand system. *The American Economic Review 70*(3), pp. 312–326.

Dixit, A. K. and J. E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review 67*(3), 297–308.

Doraszelski, U. and J. Jaumandreeu (2013). R&D and productivity: Estimating endogenous productivity. *Review of Economic Studies 80*, 1338–1383.

Dotson, J. P. and G. M. Allenby (2010). Investigating the strategic influence of customer and employee satisfaction on firm financial performance. *Marketing Science*.

Dubé, J.-P. and P. K. Chintagunta (2003). Comment on "Bayesian analysis of simultaneous demand and supply". *Quantitative Marketing and Economics 1*, 293–298.

Dunne, T., S. Klimek, and J. A. Schmitz (2009). Does foreign competition spur productivity? Evidence from post-WWII U.S. cement manufacturing. *Federal Reserve Bank of Minneapolis Staff Report*.

EPA (2009). Regulatory Impact Analysis: National Emission Standards for Hazardous Air Pollutants from the Portland Cement Manufacturing Industry.

Fabinger, M. and E. G. Weyl (2014). A tractable approach to pass-through patterns with applications to international trade. *Available at SSRN 2194855*.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *The American Economic Review*, 157–177.

Fisher, F. (1963). Uncorrelated disturbances and identifiability criteria. *International Economic Review 4*, 134–152.

Fisher, F. (1965). Near-identifiability and the variances of the disturbance terms. *Econometrica 33*, 409–419.

Fowlie, M., M. Reguant, and S. P. Ryan (2016). Market-based emissions regulation and industry dynamics. *Journal of Political Economy 124*(1), 249–302.

Gandhi, A. and J.-F. Houde (2015). Measuring substitution pattern in differentiated products industries.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*, 1029–1054.

Hausman, J. (1996). Valuation of new goods under perfect and imperfect competition. In *The Economics of New Goods, Studies in Income and Wealth* Vol 58, ed. by T. Bresnahan and R. Gordon. Chicago: National Bureau of Economic Research.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*(6), 1251–1271.

Hausman, J. A. and W. E. Taylor (1983). Identification in linear simultaneous equations models with covariance restrictions: An instrumental variables interpretation. *Econometrica 51*(5), 1527–1549.

Heckman, J. J. (1979, January). Sample Selection Bias as a Specification Error. *Econometrica 47*(1), 153.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016). Quantifying the sources of firm heterogeneity. *Quarterly Journal of Economics 131*(3), 1291–1364.

Koopmans, T. C., H. Rubin, and R. B. Leipnik (1950). *Statistical Inference in Dynamic Economic Models (Cowles Commission Monograph 10)*, Chapter Measuring the Equation Systems of Dynamic Economics. New York: John Wiley & Sons.

Leamer, E. E. (1981). Is it a demand curve, or is it a supply curve? Partial identification through inequality constraints. *The Review of Economics and Statistics 63*(3), 319–327.

Li, S., J. Mazur, Y. Park, J. Roberts, A. Sweeting, and J. Zhang (2018). Endogenous and selective service choices after airline mergers.

Macher, J., N. H. Miller, and M. Osborne (2018). Competition and technology adoption: Evidence from precalciner kilns in the cement industry.

Matzkin, R. L. (2016). On independence conditions in nonseparable models: Observable and unobservable instruments. *Journal of Econometrics 191*, 302–311.

Miller, N. H. and M. Weinberg (2017). Understanding the prices effects of the MillerCoors joint venture. *Econometrica 85*(6), 1763–1791.

National Stone Council (2008). *Limestone Quarrying and Processing: A Life-Cycle Inventory*. Report Prepared by University of Tennessee Center for Clean Products.

Nevo, A. (1998). Identification of the oligopoly solution concept in a differentiated-products industry. *Economics Letters 59*, 391–395.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Nevo, A. and A. Rosen (2012). Identification with imperfect instruments. *Review of Economics and Statistics 94*(3), 659–671.

Nevo, A. and C. Wolfram (2002). Why do manufacturers issue coupons? An empirical analysis of breakfast cereals. *RAND Journal of Economics 33*(2), 319–339.

Otter, T., T. J. Gilbride, and G. M. Allenby (2011). Testing models of strategic behavior characterized by conditional likelihoods. *Marketing Science 30*, 686–701.

Porter, R. H. (1983). A Study of Cartel Stability: The Joint Executive Committee, 1880-1886. *The Bell Journal of Economics 14*(2), 301.

Romeo, C. (2010). Filling out the instrument set in mixed logit demand systems for aggregate data. Available at SSRN: https://ssrn.com/abstract=1593590 or http://dx.doi.org/10.2139/ssrn.1593590.

Rosse, J. N. (1970, March). Estimating Cost Function Parameters Without Using Cost Data: Illustrated Methodology. *Econometrica 38*(2), 256.

Rossi, P., G. Allenby, and R. McColloch (2005). *Bayesian Statistics and Marketing*. Wiley.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica 39*, 577–592.

Schultz, H. (1928). *Statistical Laws of Demand and Supply*. The University of Chicago Press.

Thomadsen, R. (2005). The effect of ownership structure on prices in geographically differentiated industries. *RAND Journal of Economics 36*, 908–929.

Waldfogel, J. (2003). Preference externalities: An empirical study of who benefits whom in differentiated-product markets. *RAND Journal of Economics 34*(3), 557–568.

Wegge, L. (1965). Identifiability criteria for a system of equations as a whole. *The Australian Journal of Statistics 7*, 67–77.

Williams, K. R. (2017). Dynamic airline pricing and seat availability.

Working, E. J. (1927, February). What Do Statistical "Demand Curves" Show? *The Quarterly Journal of Economics 41*(2), 212.

Yang, S., Y. Chen, and G. M. Allenby (2003). Bayesian analysis of simultaneous demand and supply. *Quantitative Marketing and Economics 1*, 251–275.

A Linear Models of Supply and Demand

In this appendix we recast the monopoly model of Section 2 in terms of supply and demand, providing an alternative proof for Proposition 2 that builds explicitly on Hayashi's (2000, chapter 3) canonical textbook treatment of simultaneous equation bias. We then develop the case of perfect competition with linear demand and marginal costs, which has many similarities to monopoly and one critical difference. The model was a primary focus of previous articles addressing demand identification using covariance restrictions (e.g., Koopmans et al. (1950); Hausman and Taylor (1983); Matzkin (2016)).

A.1 Intuition from Simultaneous Equations: A Link to Hayashi

To start, given the first-order conditions of the monopolist, $p_t + (\frac{dq}{dp})^{-1} q_t = \gamma + \eta_t$ for $\frac{dq}{dp} = \beta$, equilibrium in the model can be characterized as follows:

$$\begin{aligned} q_t^d &= \alpha + \beta p_t + \xi_t && \text{(demand)} \\ q_t^s &= \beta\gamma - \beta p_t + \nu_t && \text{(supply)} \\ q_t^d &= q_t^s && \text{(equilibrium)} \end{aligned} \tag{A.1}$$

where $\nu_t \equiv \beta \eta_t$. The only distinction between this model and that of Hayashi is that slope of the supply schedule is determined (solely) by the price parameter of the demand equation, rather than by the increasing marginal cost schedules of perfect competitors.[28]

If market power is the reason that the supply schedule slopes upwards, as it is with our monopoly example, then uncorrelatedness suffices for identification because the model fully pins down how firms adjust prices with demand shocks. Repeating the steps of Hayashi, we have:

$$\beta^{OLS} \equiv plim\left(\hat{\beta}^{OLS}\right) = \beta\left(\frac{Var(\nu) - Var(\xi)}{Var(\nu) + Var(\xi)}\right) \tag{A.2}$$

If variation in the data arises solely due to cost shocks (i.e., $Var(\xi) = 0$) then the OLS estimator is consistent for β. If instead variation arises solely due to demand shocks (i.e., $Var(\nu) = 0$) then the OLS estimator is consistent for $-\beta$. A third special case arises if the demand and cost shocks have equal variance (i.e., $Var(\nu) = Var(\xi)$). Then $\beta^{OLS} = 0$, exactly halfway between the demand slope (β) and the supply slope ($-\beta$). Thus the adjustment required to bring the OLS coefficient in line with either the demand or supply slope is maximized, in terms of absolute value.

It is when variation in the data arises due both cost and demand shocks that the OLS estimate is difficult to interpret. With uncorrelatedness, however, the OLS residuals provide the information required to correct bias. A few lines of algebra obtain:

Lemma A.1. *Under uncorrelatedness, we have*

$$\beta^2 = \left(\beta^{OLS}\right)^2 + \frac{Cov(q, \xi^{OLS})}{Var(p)}. \tag{A.3}$$

[28]A implication of equation (A.1) is that it can be possible to estimate demand parameters by estimating the *supply-side* of the model, taking as given the demand system and the nature of competition. We are aware of precisely one article that employs such a method: Thomadsen (2005) estimates a model of price competition among spatially-differentiated duopolists with (importantly) constant marginal costs.

and

$$Cov(q, \xi^{OLS}) = \frac{Var(\nu)Var(\xi)}{Var(\nu) + Var(\xi)}. \tag{A.4}$$

Proof: See appendix D.

The first equation is a restatement of Proposition 2. The second equation expresses the correction term as function of $Var(\nu)$ and $Var(\xi)$. Notice that the correction term equals zero if variation in the data arises solely due to either cost or demand shocks—precisely the cases for which OLS estimator obtains β and $-\beta$, respectively. Further, the correction term is maximized if $Var(\nu) = Var(\xi)$ which, as developed above, is when the largest adjustment is required because $\beta^{OLS} = 0$.

A.2 Perfect Competition

As a point of comparison, consider perfect competition with linear demand and supply curves. The model is used elsewhere to illustrate the identifying power of covariance restrictions (e.g., Koopmans et al. (1950); Hausman and Taylor (1983); Matzkin (2016)). Let marginal costs be given by $c = x'\gamma + \lambda q + \eta$. Firms are price-takers and each has a first-order condition given by $p = x'\gamma + \lambda q + \eta$. The firm-specific supply curve is $q^s = -\frac{1}{\lambda}x'\gamma + \frac{1}{\lambda}p - \frac{\eta}{\lambda}$. Aggregating across firms and assuming with linearity in demand, we have the following market-level system of equations:

$$\begin{aligned} Q^D &= \beta p + x'\alpha + \xi && \text{(Demand)} \\ Q^S &= \frac{J}{\lambda}p - \frac{J}{\lambda}x'\gamma - \frac{J}{\lambda}\eta && \text{(Supply)} \\ Q^D &= Q^S && \text{(Equilibrium)} \end{aligned} \tag{A.5}$$

where Q^D and Q^S represent market quantity demanded and supplied, respectively. The supply slope depends on the number of firms and the slope of the marginal costs—in stark relief to the monopoly problem in which the supply slope was fully determined by the demand parameter (equation (A.1)).

In this setting, uncorrelatedness allows for the consistent estimation of the price coefficient, but only if the supply slope $\frac{J}{\lambda}$ is known. This mimics our result for oligopoly with constant marginal costs, which, in the limit of perfect competition, yields a flat supply curve. Hausman and Taylor (1983) propose the following methodology: (i) estimate the supply-schedule using an exclusion restriction $\gamma_{[k]} = 0$ for some k; (ii) recover estimates of the supply-side shock; (iii) use these estimated supply-side errors as instruments in demand estimation. Under uncorrelatedness these supply-side errors are orthogonal to demand-shock. (Though it is now understood that a method-of-moments estimator that combines uncorrelatedness with the exclusion restriction would be more efficient.) Matzkin (2016) proposes a similar procedure but relaxes the assumption of linearity.

It is possible demonstrate identification using the methods developed above for models with market power. Indeed, this can be seen as an extension of Corollary 4 because Cournot converges to perfect competition as $J \to \infty$. The OLS estimation of demand yields:

$$\beta^{OLS} \equiv plim(\hat{\beta}^{OLS}) = \beta + \frac{Cov(\xi, p^*)}{Var(p^*)}$$

Tracing the steps provided in Section 2 for the monopoly model, uncorrelatedness implies

$$Cov(\xi, Q) = Cov(\xi^{OLS}, Q) + \frac{\lambda}{J}\frac{Cov(\xi, Q)}{Var(p^*)}Cov(p^*, Q)$$

where ξ^{OLS} is a vector of OLS residuals. Solving for $Cov(\xi, Q)$ and plugging into the probability limit of the OLS estimator yields

$$\beta = \beta^{OLS} - \frac{1}{\frac{J}{\lambda} - \beta^{OLS}}Cov(\xi^{OLS}, Q) \tag{A.6}$$

It follows that β is point identified if the supply slope $\frac{J}{\lambda}$ is known. With an exclusion restriction, $\gamma_{[k]} = 0$, an estimator could be developed using equation (A.6). It would be asymptotically equivalent to the Hausman and Taylor (1983) estimator, and less efficient than the corresponding method-of-moments estimator.

B Generality of Demand

The demand system of equation (5) is sufficiently flexible to nest monopolistic competition with linear demands (e.g., as in the motivating example) and the discrete choice demand models that support much of the empirical research in industrial organization. We illustrate with some typical examples:

1. *Nested logit demand:* Following the exposition of Cardell (1997), let the firms be grouped into $g = 0, 1, \ldots, G$ mutually exclusive and exhaustive sets, and denote the set of firms in group g as $\mathscr{J}_g$. An outside good, indexed by $j = 0$, is the only member of group 0. Then the left-hand-side of equation (5) takes the form

   $$h(s_j, w_j; \sigma) \equiv \ln(s_j) - \ln(s_0) - \sigma \ln(\bar{s}_{j|g})$$

   where $\bar{s}_{j|g} = \sum_{j\in\mathscr{J}_g} \frac{s_j}{\sum_{j\in\mathscr{J}_g} s_j}$ is the market share of firm j within its group. The parameter $\sigma \in [0, 1)$ determines the extent to which consumers substitute disproportionately among firms within the same group. If $\sigma = 0$ then the logit model obtains. We can construct the markup by calculating the total derivative of h with respect to s. At the Bertrand-Nash equilibrium,

   $$\frac{dh_j}{ds_j} = \frac{1}{s_j\left(\frac{1}{1-\sigma} - s_j + \frac{\sigma}{1-\sigma}\bar{s}_{j|g}\right)}.$$

   Thus, we verify that the derivatives can be expressed as a function of data and the nonlinear parameters, allowing for three-step estimation. As we show in our second application, if uncorrelatedness is combined with a supplemental moment then the full set of parameters can be recovered (Section 5.2).

2. *Random coefficients logit demand:* Modifying slightly the notation of Berry (1994), let the

indirect utility that consumer $i = 1, \ldots, I$ receives from product j be

$$u_{ij} = \beta p_j + x_j'\alpha + \xi_j + \left[\sum_k x_{jk}\sigma_k\zeta_{ik}\right] + \epsilon_{ij}$$

where x_{jk} is the kth element of x_j, ζ_{ik} is a mean-zero consumer-specific demographic characteristic, and ϵ_{ij} is a logit error. We have suppressed market subscripts for notational simplicity. Decomposing the right-hand side of the indirect utility equation into $\delta_j = \beta p_j + x_j'\alpha + \xi_j$ and $\mu_{ij} = \sum_k x_{jk}\sigma_k\zeta_{ik}$, the probability that consumer i selects product j is given by the standard logit formula

$$s_{ij} = \frac{\exp(\delta_j + \mu_{ij})}{\sum_k \exp(\delta_k + \mu_{ik})}.$$

Integrating yields the market shares: $s_j = \frac{1}{I}\sum_i s_{ij}$. Berry et al. (1995) prove that a contraction mapping recovers, for any candidate parameter vector $\tilde{\sigma}$, the vector $\delta(s, \tilde{\sigma})$ that equates these market shares to those observed in the data. This "mean valuation" is $h(s_j, w_j; \tilde{\sigma})$ in our notation. The three-stage estimator can be applied to recover the price coefficient, again taking some $\tilde{\sigma}$ as given. At the Bertrand-Nash equilibrium, dh_j/ds_j takes the form

$$\frac{dh_j}{ds_j} = \frac{1}{\frac{1}{I}\sum_i s_{ij}(1 - s_{ij})}.$$

Thus, with the uncorrelatedness assumption the linear parameters can be recovered given the candidate parameter vector $\tilde{\sigma}$. The identification of σ is a distinct issue that has received a great deal of attention from theoretical and applied research (e.g., Waldfogel (2003); Romeo (2010); Berry and Haile (2014); Gandhi and Houde (2015); Miller and Weinberg (2017)).

3. *Constant elasticity demand:* With a substitution of $f(p_{jt})$ for p_{jt} into equation (5), the constant elasticity of substitution (CES) demand model of Dixit and Stiglitz (1977) also can be incorporated:

$$\ln(q_{jt}/q_t) = \alpha + \beta \ln\left(\frac{p_{jt}}{\Pi_t}\right) + \xi_{jt}$$

where q_t is an observed demand shifter, Π_t is a price index, and β provides the constant elasticity of demand. This model is often used in empirical research on international trade and firm productivity (e.g., De Loecker (2011); Doraszelski and Jaumandreeu (2013)). Due to the constant elasticity, profit-maximization implies $Cov(p, \xi) = 0$, and OLS produces unbiased estimates of the demand parameters. Indeed, this is an excellent illustration of our basic argument: so long as the data generating process is sufficiently well understood, it is possible to characterize the bias of OLS estimates. We opt to focus on semi-linear demand throughout this paper for analytical tractability.

Some demand systems are more difficult to reconcile with equation (5). Consider the linear demand system, $q_{jt} = \alpha_j + \sum_k \beta_{jk}p_k + \xi_{jt}$, which sometimes appears in identification proofs (e.g., Nevo (1998)) but is seldom applied empirically due to the large number of price coefficients. In principle, the system could be formulated such that $h(q_{jt}, w_{jt}; \sigma) \equiv q_{jt} - \sum_{k \neq j} \beta_{jk}p_k$ and uncorrelatedness assumptions could be used to identify the β_{jj} and α_j coefficients. This would require, however, that the econometrician have other sources of identification for the β_{jk} ($j \neq k$) coefficients, which seems unlikely. The same problem arises with the almost ideal demand system of Deaton and Muellbauer (1980).

C Two-Stage Estimation

In the presence of an additional restriction, we can produce a more precise estimator that can be calculated with one fewer stage. When the observed cost and demand shifters are uncorrelated, there is no need to project the price on demand covariates when constructing a consistent estimate, and one can proceed immediately using the OLS regression. We formalize the additional restriction and the estimator below.

Assumption 5: *Let the parameters $\alpha^{(k)}$ and $\gamma^{(k)}$ correspond to the demand and supply coefficients for covariate k in X. For any two covariates k and l, $Cov(\alpha^{(k)}x^{(k)}, \gamma^{(l)}x^{(l)}) = 0$.*

Proposition C.1. *Under assumptions 1-3 and 5, a consistent estimate of the price parameter β is given by*

$$\hat{\beta}^{\text{2-Stage}} = \frac{1}{2}\left(\hat{\beta}^{OLS} - \frac{\hat{Cov}\left(p, \frac{dh}{dq}q\right)}{\hat{Var}(p)} - \sqrt{\left(\hat{\beta}^{OLS} + \frac{\hat{Cov}\left(p, \frac{dh}{dq}q\right)}{\hat{Var}(p)}\right)^2 + 4\frac{\hat{Cov}\left(\hat{\xi}^{OLS}, \frac{dh}{dq}q\right)}{\hat{Var}(p)}}\right) \tag{C.1}$$

when the auxiliary condition, $\beta < \frac{Cov(p^*,\xi)}{Var(p^*)}\frac{Var(p)}{Var(p^*)} - \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}$, holds.

The estimator can be expressed entirely in terms of the data, the OLS coefficient, and the OLS residuals. The first stage is an OLS regression of $h(q;\cdot)$ on p and x, and the second stage is the construction of the estimator as in equation (C.1). Thus, we eliminate the step of projecting p on x. This estimator will be consistent under the assumption that any covariate affecting demand does not covary with marginal cost. The auxiliary condition parallels that of the three-stage estimator, and we expect that it will hold in typical cases.

D Proofs

Lemma: A Consistent and Unbiased Estimate for ξ

The following proof shows a consistent and unbiased estimate for the unobserved term in a linear regression when one of the covariates is endogenous. Though demonstrated in the context of semi-linear demand, the proof also applies for any endogenous covariate, including when (transformed) quantity depends on a known transformation of price, as no supply-side assumptions are required. For example, we may replace p with $\ln p$ everywhere and obtain the same results.

Lemma D.1. *A consistent and unbiased estimate of ξ is given by $\xi_1 = \xi^{OLS} + \left(\beta^{OLS} - \beta\right)p^*$*

We can construct both the true demand shock and the OLS residuals as:

$$\xi = h(q) - \beta p - x'\alpha$$
$$\xi^{OLS} = h(q) - \beta^{OLS}p - x'\alpha^{OLS}$$

where this holds even in small samples. Without loss of generality, we assume $E[\xi] = 0$. The true demand shock is given by $\xi_0 = \xi^{OLS} + (\beta^{OLS} - \beta)p + x'(\alpha^{OLS} - \alpha)$. We desire to show that an alternative estimate of the demand shock, $\xi_1 = \xi^{OLS} + (\beta^{OLS} - \beta)p^*$, is consistent and unbiased. (This eliminates the need to estimate the true α parameters). It suffices to show that $(\beta^{OLS} - \beta)p^* \to (\beta^{OLS} - \beta)p + x'(\alpha^{OLS} - \alpha)$. Consider the projection matrices

$$\begin{aligned} Q &= I - P(P'P)^{-1}P' \\ M &= I - X(X'X)^{-1}X', \end{aligned}$$

where P is an $N \times 1$ matrix of prices and X is the $N \times k$ matrix of covariates x. Denote $Y \equiv h(q) = P\beta + X\alpha + \xi$. Our OLS estimators can be constructed by a residualized regression

$$\begin{aligned} \alpha^{OLS} &= \left((XQ)'QX\right)^{-1}(XQ)'Y \\ \beta^{OLS} &= \left((PM)'MP\right)^{-1}(PM)'Y. \end{aligned}$$

Therefore

$$\begin{aligned} \alpha^{OLS} =& \left(X'QX\right)^{-1}\left(X'QP\beta + X'QX\alpha + X'Q\xi\right) \\ =& \alpha + \left(X'QX\right)^{-1}X'Q\xi. \end{aligned}$$

Similarly,

$$\begin{aligned} \beta^{OLS} =& \left(P'MP\right)^{-1}\left(P'MP\beta + P'MX\alpha + P'M\xi\right) \\ =& \beta + \left(P'MP\right)^{-1}P'M\xi. \end{aligned}$$

We desire to show

$$P^*(\beta^{OLS} - \beta) \to P(\beta^{OLS} - \beta) + X(\alpha^{OLS} - \alpha).$$

Note that $P^* = MP$. Then

$$\begin{aligned} P^*(\beta^{OLS} - \beta) \to& P(\beta^{OLS} - \beta) + X(\alpha^{OLS} - \alpha) \\ MP\left(P'MP\right)^{-1}P'M\xi \to& P\left(P'MP\right)^{-1}P'M\xi + X\left(X'QX\right)^{-1}X'Q\xi \\ -X(X'X)^{-1}X'P\left(P'MP\right)^{-1}P'M\xi \to& X\left(X'QX\right)^{-1}X'Q\xi \\ -X(X'X)^{-1}X'P\left(P'MP\right)^{-1}P'\left[I - X(X'X)^{-1}X'\right]\xi \to& X\left(X'QX\right)^{-1}X'\left[I - P(P'P)^{-1}P'\right]\xi \\ -X(X'X)^{-1}X'P\left(P'MP\right)^{-1}P'\xi \to& X\left(X'QX\right)^{-1}X'\xi \\ +X(X'X)^{-1}X'P\left(P'MP\right)^{-1}P'X(X'X)^{-1}X'\xi \quad & - X\left(X'QX\right)^{-1}X'P(P'P)^{-1}P'\xi. \end{aligned}$$

We will show that the following two relations hold, which proves consistency and completes the proof.

1. $X(X'X)^{-1}X'P\left(P'MP\right)^{-1}P'\xi = X\left(X'QX\right)^{-1}X'P(P'P)^{-1}P'\xi$
2. $X(X'X)^{-1}X'P\left(P'MP\right)^{-1}P'X(XX\left(X'QX\right)^{-1}X'X)^{-1}X'\xi \to X\left(X'QX\right)^{-1}X'\xi$

Part 1: Equivalence

It suffices to show that $X(X'X)^{-1}X'P\,(P'MP)^{-1} = X\,(X'QX)^{-1}\,X'P(P'P)^{-1}$.

$$
\begin{aligned}
X(X'X)^{-1}X'P\left(P'MP\right)^{-1} =& X\left(X'QX\right)^{-1}X'P(P'P)^{-1} \\
X(X'X)^{-1}X'P =& X\left(X'QX\right)^{-1}X'P(P'P)^{-1}\left(P'MP\right) \\
X(X'X)^{-1}X'P =& X\left(X'QX\right)^{-1}X'P(P'P)^{-1}\left(P'P\right) \\
&- X\left(X'QX\right)^{-1}X'P(P'P)^{-1}\left(P'X(X'X)^{-1}X'P\right) \\
X(X'X)^{-1}X'P =& X\left(X'QX\right)^{-1}X'P \\
&- X\left(X'QX\right)^{-1}X'\left[I-Q\right]X(X'X)^{-1}X'P \\
X(X'X)^{-1}X'P =& X\left(X'QX\right)^{-1}X'P \\
&- X\left(X'QX\right)^{-1}X'X(X'X)^{-1}X'P \\
&+ X\left(X'QX\right)^{-1}X'QX(X'X)^{-1}X'P \\
X(X'X)^{-1}X'P =& X(X'X)^{-1}X'P
\end{aligned}
$$

QED.

Part 2: Consistency (and Unbiasedness)

Because $X(X'X)^{-1}X'P = X\,(X'QX)^{-1}\,X'P(P'P)^{-1}\,(P'MP)$, as shown above:

$$
\begin{aligned}
X(X'X)^{-1}X'P\left(P'MP\right)^{-1}P'X(X'X)^{-1}X'\xi &\to X\left(X'QX\right)^{-1}X'\xi \\
X\left(X'QX\right)^{-1}X'P(P'P)^{-1}P'X(X'X)^{-1}X'\xi &\to X\left(X'QX\right)^{-1}X'\xi \\
X\left(X'QX\right)^{-1}X'\left[I-Q\right]X(X'X)^{-1}X'\xi &\to X\left(X'QX\right)^{-1}X'\xi \\
X\left(X'QX\right)^{-1}X'X(X'X)^{-1}X'\xi &\to X\left(X'QX\right)^{-1}X'\xi \\
-X(X'X)^{-1}X'\xi & \\
X\left(X'QX\right)^{-1}X'\xi - X(X'X)^{-1}X'\xi &\to X\left(X'QX\right)^{-1}X'\xi \\
X(X'X)^{-1}X'\xi &\to 0.
\end{aligned}
$$

The last line, where the projection of ξ onto the exogenous covariates X converges to zero, holds by assumption. We say that two vectors converge if the mean absolute deviation goes to zero as the sample size gets large. Note that also $E[X(X'X)^{-1}X'\xi] = 0$, so ξ_1 is both a consistent and unbiased estimate for ξ_0. QED.

Proof of Proposition 3 (Set Identification)

From the text, we have $\hat{\beta}^{OLS} \xrightarrow{p} \beta + \frac{Cov(p^*,\xi)}{Var(p^*)}$. The general form for a firm's first-order condition is $p = c + \mu$, where c is the marginal cost and μ is the markup. We can write $p = p^* + \hat{p}$, where $\hat{p}$ is the projection of p onto the exogenous demand variables, X. By assumption, $c = X\gamma + \eta$. If we substitute the first-order condition $p^* = X\gamma + \eta + \mu - \hat{p}$ into the bias term from the OLS regression, we obtain

$$\begin{aligned}\frac{Cov(p^*,\xi)}{Var(p^*)} &= \frac{Cov(\xi, X\gamma+\eta+\mu-\hat{p})}{Var(p^*)}\\ &= \frac{Cov(\xi,\eta)}{Var(p^*)}+\frac{Cov(\xi,\mu)}{Var(p^*)}\end{aligned}$$

where the second line follows from the exogeneity assumption ($E[X\xi]=0$). Under our demand assumption, the unobserved demand shock may be written as $\xi = h(q) - x\alpha - \beta p$. At the probability limit of the OLS estimator, we can construct a consistent estimate of the unobserved demand shock as $\xi = \xi^{OLS} + \left(\beta^{OLS}-\beta\right)p^*$ (see Lemma D.1 above). From the prior step in this proof, $\beta^{OLS}-\beta = \frac{Cov(\xi,\eta)}{Var(p^*)}+\frac{Cov(\xi,\mu)}{Var(p^*)}$. Therefore, $\xi = \xi^{OLS} + \left(\frac{Cov(\eta,\xi)}{Var(p^*)}+\frac{Cov(\mu,\xi)}{Var(p^*)}\right)p^*$. This implies

$$\begin{aligned}\frac{Cov\left(\xi,\mu\right)}{Var(p^*)} &= \frac{Cov\left(\xi^{OLS},\mu\right)}{Var(p^*)}+\left(\frac{Cov(\xi,\eta)}{Var(p^*)}+\frac{Cov\left(\xi,\mu\right)}{Var(p^*)}\right)\frac{Cov(p^*,\mu)}{Var(p^*)}\\ \frac{Cov\left(\xi,\mu\right)}{Var(p^*)}\left(1-\frac{Cov(p^*,\mu)}{Var(p^*)}\right) &= \frac{Cov\left(\xi^{OLS},\mu\right)}{Var(p^*)}+\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\mu)}{Var(p^*)}\\ \frac{Cov\left(\xi,\mu\right)}{Var(p^*)} &= \frac{1}{1-\frac{Cov(p^*,\mu)}{Var(p^*)}}\frac{Cov\left(\xi^{OLS},\mu\right)}{Var(p^*)}+\frac{1}{1-\frac{Cov(p^*,\mu)}{Var(p^*)}}\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\mu)}{Var(p^*)}\end{aligned}$$

When we substitute this expression in for β^{OLS}, we obtain

$$\begin{aligned}\beta^{OLS} &= \beta+\frac{Cov(\xi,\eta)}{Var(p^*)}+\frac{1}{1-\frac{Cov(p^*,\mu)}{Var(p^*)}}\frac{Cov\left(\xi^{OLS},\mu\right)}{Var(p^*)}+\frac{\frac{Cov(p^*,\mu)}{Var(p^*)}}{1-\frac{Cov(p^*,\mu)}{Var(p^*)}}\frac{Cov(\xi,\eta)}{Var(p^*)}\\ \beta^{OLS} &= \beta+\frac{1}{1-\frac{Cov(p^*,\mu)}{Var(p^*)}}\frac{Cov\left(\xi^{OLS},\mu\right)}{Var(p^*)}+\frac{1}{1-\frac{Cov(p^*,\mu)}{Var(p^*)}}\frac{Cov(\xi,\eta)}{Var(p^*)}\end{aligned}$$

Thus, we obtain an expression for the OLS estimator in terms of the OLS residuals, the residualized prices, the markup, and the correlation between unobserved demand and cost shocks. If the markup can be parameterized in terms of observables and the correlation in unobserved shocks can be calibrated, we have a method to estimate β from the OLS regression. Under our supply and demand assumptions, $\mu = -\frac{1}{\beta}\frac{dh}{dq}q$, and plugging in obtains the first equation of the proposition:

$$\beta^{OLS} = \beta - \frac{1}{\beta+\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}}\frac{Cov\left(\xi^{OLS},\frac{dh}{dq}q\right)}{Var(p^*)}+\beta\frac{1}{\beta+\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}}\frac{Cov(\xi,\eta)}{Var(p^*)}.$$

The second equation in the proposition is obtained by rearranging terms. QED.

Proof of Proposition 4 (Point Identification)

Part (1). We first prove the sufficient condition, i.e., that under assumptions 1 and 2, β is the lower root of equation (11) if the following condition holds:

$$0 \leq \beta^{OLS} \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} \tag{D.1}$$

Consider a generic quadratic, $ax^2 + bx + c$. The roots of the quadratic are $\frac{1}{2a}\left(-b \pm \sqrt{b^2 - 4ac}\right)$. Thus, if $4ac < 0$ and $a > 0$ then the upper root is positive and the lower root is negative. In equation (11), $a = 1$, and $4ac < 0$ if and only if equation (D.1) holds. Because the upper root is positive, $\beta < 0$ must be the lower root, and point identification is achieved given knowledge of $Cov(\xi, \eta)$. QED.

Part (2). In order to prove the necessary and sufficient condition for point identification, we first state and prove a lemma:

Lemma D.2. *The roots of equation* (11) *are* β *and* $\frac{Cov(p^*,\xi)}{Var(p^*)} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}$.

Proof of Lemma D.2. We first provide equation (11) for reference:

$$\begin{aligned} 0 &= \beta^2 \\ &+ \left(\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov(\xi, \eta)}{Var(p^*)} - \beta^{OLS}\right)\beta \\ &+ \left(-\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\right) \end{aligned}$$

To find the roots, begin by applying the quadratic formula

$$
\begin{aligned}
(r_1, r_2) &= \frac{1}{2}\left(-B \pm \sqrt{B^2 - 4AC}\right) \\
&= \frac{1}{2}\Bigg(\beta^{OLS} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)} \\
&\quad \pm \sqrt{\left(\beta^{OLS} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)}\right)^2 + 4\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + 4\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}}\Bigg) \\
&= \frac{1}{2}\Bigg[\beta^{OLS} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)} \\
&\quad \pm \Bigg(\left(\beta^{OLS} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + \left(\frac{Cov(\xi, \eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\xi, \eta)}{Var(p^*)}\left(\beta^{OLS} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right) \\
&\quad + 4\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + 4\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\Bigg)^{\frac{1}{2}}\Bigg] \\
&= \frac{1}{2}\Bigg(\beta^{OLS} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)} \\
&\quad \pm \sqrt{\left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} + \left(\frac{Cov(\xi, \eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\xi, \eta)}{Var(p^*)}\left(\beta^{OLS} - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)}\Bigg)
\end{aligned}
\tag{D.2}
$$

Looking inside the radical, consider the first part: $\left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}$

$$
\begin{aligned}
&\left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} \\
&= \left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\xi - p^*(\beta^{OLS} - \beta), \frac{dh}{dq}q\right)}{Var(p^*)} \\
&= \left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)} - 4\frac{Cov(p^*, \xi)}{Var(p^*)}\frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)} \\
&= \left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)} - 4\left(\frac{Cov(\xi, \eta)}{Var(p^*)} + \frac{Cov(\xi, -\frac{1}{\beta}\frac{dh}{dq}q)}{Var(p^*)}\right)\frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)} \\
&= \left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)}\left(1 + \frac{1}{\beta}\frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}\right) - 4\frac{Cov(\xi, \eta)}{Var(p^*)}\frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}
\end{aligned}
\tag{D.3}
$$

To simplify this expression, it is helpful to use the general form for a firm's first-order condition, $p = c + \mu$, where c is the marginal cost and μ is the markup. We can write $p = p^* + \hat{p}$, where $\hat{p}$ is the projection of p onto the exogenous demand variables, X. By assumption, $c = X\gamma + \eta$. It follows that

$$
\begin{aligned}
p^* &= X\gamma + \eta + \mu - \hat{p} \\
&= X\gamma + \eta - \frac{1}{\beta}\frac{dh}{dq}q - \hat{p}
\end{aligned}
$$

Therefore

$$Cov(p^*, \xi) = Cov(\xi, \eta) - \frac{1}{\beta} Cov(\xi, \frac{dh}{dq} q)$$

and

$$Cov(\xi, \frac{dh}{dq} q) = -\beta \left(Cov(p^*, \xi) - Cov(\xi, \eta) \right)$$

$$\frac{Cov(\xi, \frac{dh}{dq} q)}{Var(p^*)} = -\beta \left(\frac{Cov(p^*, \xi)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)} \right) \tag{D.4}$$

Returning to equation (D.3), we can substitute using equation (D.4) and simplify:

$$\begin{aligned}
&\left(\beta^{OLS} + \frac{Cov(p^*, \frac{dh}{dq} q)}{Var(p^*)} \right)^2 + 4 \frac{Cov\left(\xi, \frac{dh}{dq} q\right)}{Var(p^*)} \left(1 + \frac{1}{\beta} \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \right) - 4 \frac{Cov(\xi, \eta)}{Var(p^*)} \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \\
= &\left(\beta^{OLS} \right)^2 + \left(\frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \right)^2 + 2\beta^{OLS} \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} - 4 \frac{Cov(\xi, \eta)}{Var(p^*)} \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \\
&+ 4 \frac{Cov\left(\xi, \frac{dh}{dq} q\right)}{Var(p^*)} + 4 \frac{1}{\beta} \frac{Cov\left(\xi, \frac{dh}{dq} q\right)}{Var(p^*)} \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \\
= &\left(\beta + \frac{Cov\left(p^*, \xi\right)}{Var(p^*)} \right)^2 + \left(\frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \right)^2 + 2 \left(\beta + \frac{Cov\left(p^*, \xi\right)}{Var(p^*)} \right) \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} - 4 \frac{Cov(\xi, \eta)}{Var(p^*)} \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \\
&- 4\beta \left(\frac{Cov(p^*, \xi)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)} \right) - 4 \left(\frac{Cov(p^*, \xi)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)} \right) \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \\
= &\left(\beta + \frac{Cov\left(p^*, \xi\right)}{Var(p^*)} \right)^2 + \left(\frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \right)^2 + 2 \left(\beta + \frac{Cov\left(p^*, \xi\right)}{Var(p^*)} \right) \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \\
&- 4\beta \left(\frac{Cov(p^*, \xi)}{Var(p^*)} \right) - 4 \left(\frac{Cov(p^*, \xi)}{Var(p^*)} \right) \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} + 4\beta \frac{Cov(\xi, \eta)}{Var(p^*)} \\
= &\beta^2 + \left(\frac{Cov\left(p^*, \xi\right)}{Var(p^*)} \right)^2 + \left(\frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \right)^2 + 2\beta \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \\
&- 2\beta \frac{Cov\left(p^*, \xi\right)}{Var(p^*)} - 2 \frac{Cov\left(p^*, \xi\right)}{Var(p^*)} \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} + 4\beta \frac{Cov(\xi, \eta)}{Var(p^*)} \\
= &\left(\left(\beta + \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)} \right) - \frac{Cov\left(p^*, \xi\right)}{Var(p^*)} \right)^2 + 4\beta \frac{Cov(\xi, \eta)}{Var(p^*)}
\end{aligned}$$

Now, consider the second part inside of the radical in equation (D.2):

$$\begin{aligned}
&\left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\left(\beta^{OLS} - \frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}\right) \\
&= \left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\left(\beta + \frac{Cov(\xi,\eta)}{Var(p^*)} - \frac{1}{\beta}\frac{Cov(\xi,\frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}\right) \\
&= \left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 - 2\beta\frac{Cov(\xi,\eta)}{Var(p^*)} - 2\left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 + 2\frac{1}{\beta}\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(\xi,\frac{dh}{dq}q)}{Var(p^*)} + 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} \\
&= -\left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 - 2\beta\frac{Cov(\xi,\eta)}{Var(p^*)} - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\left(\frac{Cov(p^*,\xi)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}\right) + 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} \\
&= \left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\beta - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\xi)}{Var(p^*)} + 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}
\end{aligned}$$

Combining yields a simpler expression for the terms inside the radical of equation (D.2):

$$\begin{aligned}
&\left(\left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)}\right) - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)}\right)^2 + 4\beta\frac{Cov(\xi,\eta)}{Var(p^*)} \\
&\quad + \left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\beta - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\xi)}{Var(p^*)} + 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} \\
&= \left(\left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)}\right) - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)}\right)^2 + \left(\frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2 \\
&\quad + 2\beta\frac{Cov(\xi,\eta)}{Var(p^*)} - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\xi)}{Var(p^*)} + 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} \\
&= \left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2
\end{aligned}$$

Plugging this back into equation (D.2), we have:

$$\begin{aligned}
(r_1, r_2) \quad &= \quad \frac{1}{2}\left(\beta^{OLS} - \frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}\right. \\
&\qquad \left. \pm\sqrt{\left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2}\right) \\
&= \quad \frac{1}{2}\left(\beta + \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} - \frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}\right. \\
&\qquad \left. \pm\sqrt{\left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2}\right)
\end{aligned}$$

The roots are given by

$$\frac{1}{2}\left(\beta+\frac{Cov\left(p^{*},\xi\right)}{Var(p^{*})}-\frac{Cov(p^{*},\frac{dh}{dq}q)}{Var(p^{*})}-\frac{Cov(\xi,\eta)}{Var(p^{*})}+\beta+\frac{Cov\left(p^{*},\frac{dh}{dq}q\right)}{Var(p^{*})}-\frac{Cov\left(p^{*},\xi\right)}{Var(p^{*})}+\frac{Cov(\xi,\eta)}{Var(p^{*})}\right)$$

$$=\beta$$

and

$$\frac{1}{2}\left(\beta+\frac{Cov\left(p^{*},\xi\right)}{Var(p^{*})}-\frac{Cov(p^{*},\frac{dh}{dq}q)}{Var(p^{*})}-\frac{Cov(\xi,\eta)}{Var(p^{*})}-\beta-\frac{Cov\left(p^{*},\frac{dh}{dq}q\right)}{Var(p^{*})}+\frac{Cov\left(p^{*},\xi\right)}{Var(p^{*})}-\frac{Cov(\xi,\eta)}{Var(p^{*})}\right)$$

$$=\frac{Cov\left(p^{*},\xi\right)}{Var(p^{*})}-\frac{Cov(p^{*},\frac{dh}{dq}q)}{Var(p^{*})}-\frac{Cov(\xi,\eta)}{Var(p^{*})}$$

which completes the proof of the intermediate result. QED.

Part (3). Consider the roots of equation (11), β and $\frac{Cov(p^{*},\xi)}{Var(p^{*})}-\frac{Cov(p^{*},\frac{dh}{dq}q)}{Var(p^{*})}-\frac{Cov(\xi,\eta)}{Var(p^{*})}$. The price parameter β may or may not be the lower root.[29] However, β is the lower root iff

$$\begin{aligned}
\beta &< \frac{Cov(p^*,\xi)}{Var(p^*)} - \frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)} \\
\beta &< -\beta\frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} + \beta\frac{Cov(p^*,-\frac{1}{\beta}\frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)} \\
\beta &< -\beta\frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} + \beta\frac{Cov(p^*,p^*-c)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)} \\
\beta &< \beta\frac{Var(p^*)}{Var(p^*)} - \beta\frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} - \beta\frac{Cov(p^*,\eta)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)} \\
0 &< -\beta\frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} - \beta\frac{Cov(p^*,\eta)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)} \\
0 &< \frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} + \frac{Cov(p^*,\eta)}{Var(p^*)} + \frac{1}{\beta}\frac{Cov(\xi,\eta)}{Var(p^*)}
\end{aligned}$$

The third line relies on the expression for the markup, $p - c = -\frac{1}{\beta}\frac{dh}{dq}q$. The final line holds because $\beta < 0$ so $-\beta > 0$. It follows that β is the lower root of (11) iff

$$-\frac{1}{\beta}\frac{Cov(\xi,\eta)}{Var(p^*)} \leq \frac{Cov\left(p^*,-\frac{1}{\beta}\xi\right)}{Var(p^*)} + \frac{Cov\left(p^*,\eta\right)}{Var(p^*)}$$

in which case β is point identified given knowledge of $Cov(\xi,\eta)$. QED.

Proof of Lemma 1 (Monotonicity in $Cov(\xi,\eta)$)

We return to the quadratic formula for the proof. The lower root of a quadratic ax^2+bx+c is $L \equiv \frac{1}{2}\left(-b-\sqrt{b^2-4ac}\right)$. In our case, $a=1$.

[29] Consider that the first root is the upper root if

$$\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)} > 0$$

because, in that case,

$$\sqrt{\left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2} = \beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)}$$

When $\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)} < 0$, then $\sqrt{\left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)}\right)^2} = -\left(\beta + \frac{Cov\left(p^*,\frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(p^*,\xi\right)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)}\right)$, and the first root is then the lower root (i.e., minus the negative value).

We wish to show that $\frac{\partial L}{\partial \gamma} < 0$, where $\gamma = Cov(\xi, \eta)$. We evaluate the derivative to obtain

$$\frac{\partial L}{\partial \gamma} = -\frac{1}{2}\left(1 + \frac{b}{(b^2 - 4c)^{\frac{1}{2}}}\right)\frac{\partial b}{\partial \gamma}.$$

We observe that, in our setting, $\frac{\partial b}{\partial \gamma} = \frac{1}{Var(p^*)}$ is always positive. Therefore, it suffices to show that

$$1 + \frac{b}{(b^2 - 4c)^{\frac{1}{2}}} > 0. \tag{D.5}$$

We have two cases. First, when $c < 0$, we know that $\left|\frac{b}{(b^2-4c)^{\frac{1}{2}}}\right| < 1$, which satisfies (D.5). Second, when $c > 0$, it must be the case that $b > 0$ also. Otherwise, both roots will be positive, invalidating the model. When $b > 0$, it is evident that the left-hand side of (D.5) is positive. This demonstrates monotonicity.

Finally, we obtain the range of values for L by examining the limits as $\gamma \to \infty$ and $\gamma \to -\infty$. From the expression for L and the result that $\frac{\partial b}{\partial \gamma}$ is a constant, we obtain

$$\lim_{\gamma \to -\infty} L = 0$$
$$\lim_{\gamma \to \infty} L = -\infty$$

When $c < 0$, the domain of the quadratic function is $(-\infty, \infty)$, which, along with monotonicity, implies the range for L of $(0, -\infty)$. When $c > 0$, the domain is not defined on the interval $(-2\sqrt{c}, 2\sqrt{c})$, but L is equal in value at the boundaries of the domain. QED.

Additionally, we note that the upper root, $U \equiv \frac{1}{2}\left(-b + \sqrt{b^2 - 4ac}\right)$ will be increasing in γ. When the upper root is a valid solution (i.e., negative), it must be the case that $c > 0$ and $b > 0$, and it is straightforward to follow the above arguments to show that $\frac{\partial U}{\partial \gamma} > 0$ and that the range of the upper root is $[-\frac{1}{2}b, 0)$.

Proof of Proposition 5 (Prior-Free Bounds)

The proof is again an application of the quadratic formula. Any generic quadratic, $ax^2 + bx + c$, with roots $\frac{1}{2}\left(-b \pm \sqrt{b^2 - 4ac}\right)$, admits a real solution if and only if $b^2 > 4ac$. Given the formulation of (11), real solutions satisfy the condition:

$$\left(\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov(\xi, \eta)}{Var(p^*)} - \beta^{OLS}\right)^2 > 4\left(-\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\right).$$

As $a = 1$, a solution is always possible if $c < 0$. This is the sufficient condition for point identification from the text. If $c > 0$, it must be the case that $b > 0$; otherwise, both roots will be positive. Therefore, a real solution is obtained if and only if $b > 2\sqrt{c}$, that is

$$\left(\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} + \frac{Cov(\xi,\eta)}{Var(p^*)} - \beta^{OLS}\right) > 2\sqrt{\left(-\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\right)}.$$

Solving for $Cov(\xi,\eta)$, we obtain the prior-free bound,

$$Cov(\xi,\eta) > Var(p^*)\beta^{OLS} - Cov(p^*, \frac{dh}{dq}q) + 2Var(p^*)\sqrt{\left(-\beta^{OLS}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\right)}.$$

This bound exists if the expression inside the radicals is positive, which is the case if and only if the sufficient condition for point identification from Proposition 4 fails. QED.

Proof of Proposition 6 (Non-Constant Marginal Costs)

Under the semi-linear marginal cost schedule of equation (14), the plim of the OLS estimator is equal to

$$\text{plim}\hat{\beta}^{OLS} = \beta + \frac{Cov(\xi, g(q))}{Var(p^*)} - \frac{1}{\beta}\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)}.$$

This is obtain directly by plugging in the first–order condition for p: $Cov(p^*,\xi) = Cov(g(q) + \eta - \frac{1}{\beta}\frac{dh}{dq}q - \hat{p}, \xi) = Cov(\xi, g(q)) - \frac{1}{\beta}Cov(\xi, \frac{dh}{dq}q)$under the assumptions. Next, we re-express the terms including the unobserved demand shocks in in terms of OLS residuals. The unobserved demand shock may be written as $\xi = h(q) - x\beta_x - \beta p$. The estimated residuals are given by $\xi^{OLS} = \xi + \left(\beta - \beta^{OLS}\right)p^*$. As $\beta - \beta^{OLS} = \frac{1}{\beta}\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov(\xi,g(q))}{Var(p^*)}$, we obtain $\xi^{OLS} = \xi + \left(\frac{1}{\beta}\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov(\xi,g(q))}{Var(p^*)}\right)p^*$. This implies

$$\begin{aligned}
Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right) &= \left(1 + \frac{1}{\beta}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\right)Cov(\xi, \frac{dh}{dq}q) - \frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}Cov(\xi, g(q)) \\
Cov\left(\xi^{OLS}, g(q)\right) &= \frac{1}{\beta}\frac{Cov(p^*, g(q))}{Var(p^*)}Cov\left(\xi, \frac{dh}{dq}q\right) + \left(1 - \frac{Cov(p^*, g(q))}{Var(p^*)}\right)Cov(\xi, g(q))
\end{aligned}$$

We write the system of equations in matrix form and invert to solve for the covariance terms that include the unobserved demand shock:

$$\begin{bmatrix} Cov(\xi, \frac{dh}{dq}q) \\ Cov(\xi, g(q)) \end{bmatrix} = \begin{bmatrix} 1 + \frac{1}{\beta}\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} & -\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} \\ \frac{1}{\beta}\frac{Cov(p^*,g(q))}{Var(p^*)} & 1 - \frac{Cov(p^*,g(q))}{Var(p^*)} \end{bmatrix}^{-1} \begin{bmatrix} Cov(\xi^{OLS}, \frac{dh}{dq}q) \\ Cov(\xi^{OLS}, g(q)) \end{bmatrix}$$

where

$$\begin{bmatrix} 1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} & -\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} \\ \frac{1}{\beta}\frac{Cov(p^*,g(q))}{Var(p^*)} & 1-\frac{Cov(p^*,g(q))}{Var(p^*)} \end{bmatrix}^{-1} =$$

$$\frac{1}{1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}-\frac{Cov(p^*,g(q))}{Var(p^*)}}\begin{bmatrix} 1-\frac{Cov(p^*,g(q))}{Var(p^*)} & \frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} \\ -\frac{1}{\beta}\frac{Cov(p^*,g(q))}{Var(p^*)} & 1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)} \end{bmatrix}.$$

Therefore, we obtain the relations

$$Cov(\xi,\frac{dh}{dq}q) = \frac{\left(1-\frac{Cov(p^*,g(q))}{Var(p^*)}\right)Cov(\xi^{OLS},\frac{dh}{dq}q)+\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}Cov(\xi^{OLS},g(q))}{1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}-\frac{Cov(p^*,g(q))}{Var(p^*)}}$$

$$Cov(\xi,g(q)) = \frac{-\frac{1}{\beta}\frac{Cov(p^*,g(q))}{Var(p^*)}Cov(\xi^{OLS},\frac{dh}{dq}q)+\left(1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}\right)Cov(\xi^{OLS},g(q))}{1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}-\frac{Cov(p^*,g(q))}{Var(p^*)}}.$$

In terms of observables, we can substitute in for $Cov(\xi,g(q))-\frac{1}{\beta}Cov\left(\xi,\frac{dh}{dq}q\right)$ in the plim of the OLS estimator and simplify:

$$\begin{aligned}
&\left(1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}-\frac{Cov(p^*,g(q))}{Var(p^*)}\right)\left(Cov(\xi,g(q))-\frac{1}{\beta}Cov\left(\xi,\frac{dh}{dq}q\right)\right)\\
=&-\frac{1}{\beta}\frac{Cov(p^*,g(q))}{Var(p^*)}Cov(\xi^{OLS},\frac{dh}{dq}q)+\left(1+\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}\right)Cov(\xi^{OLS},g(q))\\
&-\frac{1}{\beta}\left(1-\frac{Cov(p^*,g(q))}{Var(p^*)}\right)Cov(\xi^{OLS},\frac{dh}{dq}q)-\frac{1}{\beta}\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}Cov(\xi^{OLS},g(q))\\
=&Cov(\xi^{OLS},g(q))-\frac{1}{\beta}Cov(\xi^{OLS},\frac{dh}{dq}q).
\end{aligned}$$

Thus, we obtain an expression for the probability limit of the OLS estimator,

$$\text{plim}\hat{\beta}^{OLS} = \beta - \frac{\frac{Cov(\xi^{OLS},\frac{dh}{dq}q)}{Var(p^*)}-\beta\frac{Cov(\xi^{OLS},g(q))}{Var(p^*)}}{\beta+\frac{Cov(p^*,\frac{dh}{dq}q)}{Var(p^*)}-\beta\frac{Cov(p^*,g(q))}{Var(p^*)}},$$

and the following quadratic β.

$$\begin{aligned}
0 = &\left(1 - \frac{Cov(p^*, g(q))}{Var(p^*)}\right)\beta^2 \\
&+ \left(\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)} - \hat{\beta}^{OLS} + \frac{Cov(p^*, g(q))}{Var(p^*)}\hat{\beta}^{OLS} + \frac{Cov(\xi^{OLS}, g(q))}{Var(p^*)}\right)\beta \\
&+ \left(-\frac{Cov(p^*, \frac{dh}{dq}q)}{Var(p^*)}\hat{\beta}^{OLS} - \frac{Cov(\xi^{OLS}, \frac{dh}{dq}q)}{Var(p^*)}\right).
\end{aligned}$$

QED.

Proof of Lemma A.1

The proof is by construction. Note that model has the solutions $p_t^* = \frac{1}{2}\left(-\frac{\alpha}{\beta} - \frac{\xi_t}{\beta} + \gamma + \frac{\nu_t}{\beta}\right)$ and $q_t^* = \frac{1}{2}(\alpha + \xi_t + \beta\gamma + \nu_t)$, where again $\nu_t \equiv \beta\eta_t$. The following objects are easily derived:

$$Cov(p, \xi) = -\frac{1}{2\beta}Var(\xi) \qquad Cov(p, \nu) = \frac{1}{2\beta}Var(\nu)2$$
$$Var(p) = \frac{Var(\nu) + Var(\xi)}{(2\beta)^2} \qquad Var(q) = \frac{1}{4}(Var(\xi) + Var(\nu))$$

Using the above, we have

$$\begin{aligned}
Cov(p, q) &= Cov(p, \alpha + \beta p + \xi) = \beta Var(p) + Cov(p, \xi) = \beta\frac{Var(\nu) + Var(\xi)}{(2\beta)^2} - \frac{2\beta}{(2\beta)^2}Var(\xi) \\
&= \frac{\beta Var(\nu) + \beta Var(\xi) - 2\beta Var(\xi)}{(2\beta)^2} = \beta\frac{Var(\nu) - Var(\xi)}{(2\beta)^2}
\end{aligned}$$

And that obtains equation (A.2):

$$plim\left(\hat{\beta}^{OLS}\right) \equiv \beta^{OLS} = \frac{Cov(p, q)}{Var(p)} = \beta\frac{Var(\nu) - Var(\xi)}{Var(\nu) + Var(\xi)}$$

Equation (A.4) requires an expression for $Cov(q, \xi^{OLS})$. Define

$$plim(\hat{\xi}^{OLS}) \equiv \xi^{OLS} = q - \alpha^{OLS} - \beta^{OLS}p$$

Then, plugging into $Cov(q, \xi^{OLS})$ using the objects derived above, we have

$$\begin{aligned}
Cov(q, \xi^{OLS}) &= Cov(q, q - \beta^{OLS} p) \\
&= Var(q) - \beta^{OLS} Cov(p, q) \\
&= \frac{1}{4}(Var(\xi) + Var(\nu)) - \left(\beta \frac{Var(\nu) - Var(\xi)}{Var(\nu) + Var(\xi)}\right)\left(\beta \frac{(Var(\nu) - Var(\xi))}{(2\beta)^2}\right) \\
&= \frac{1}{4}\left(\frac{[Var(\xi) + Var(\nu)]^2 - [Var(\nu) - Var(\xi)]^2}{Var(\nu) + Var(\xi)}\right) \\
&= \frac{Var(\xi) Var(\nu)}{Var(\nu) + Var(\xi)}
\end{aligned}$$

We turn now to equation (A.3). Based on the above, we have that

$$\frac{Cov(q, \xi^{OLS})}{Var(p)} = \left(\frac{Var(\xi)Var(\nu)}{Var(\nu) + Var(\xi)}\right) \frac{(2\beta)^2}{Var(\nu) + Var(\xi)} = (2\beta)^2 \frac{Var(\xi)Var(\nu)}{[Var(\nu) + Var(\xi)]^2}$$

and now only few more lines of algebra are required:

$$\begin{aligned}
(\beta^{OLS})^2 + \frac{Cov(q, \xi^{OLS})}{Var(p)} &= \beta^2 \left[\frac{Var(\nu) - Var(\xi)}{Var(\nu) + Var(\xi)}\right]^2 + (2\beta)^2 \frac{Var(\xi)Var(\nu)}{[Var(\nu) + Var(\xi)]^2} \\
&= \frac{\beta^2 [Var^2(\nu) + Var^2(\xi) - 2Var(\nu)Var(\xi)] + 4\beta^2 Var(\nu)Var(\xi)}{[Var(\nu) + Var(\xi)]^2} \\
&= \frac{\beta^2 [Var^2(\nu) + Var^2(\xi) + 2Var(\nu)Var(\xi)]}{[Var(\nu) + Var(\xi)]^2} \\
&= \beta^2 \frac{[Var(\nu) + Var(\xi)]^2}{[Var(\nu) + Var(\xi)]^2} = \beta^2
\end{aligned}$$

QED.

Proof of Proposition C.1 (Two-Stage Estimator)

Suppose that, in addition to assumptions 1-3, that marginal costs are uncorrelated with the exogenous demand factors (Assumption 5). Then, the expression $\frac{1}{\beta + \frac{Cov\left(p, \frac{dh}{dq} q\right)}{Var(p)}} \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq} q\right)}{Var(p)}$ is equal to $\frac{1}{\beta + \frac{Cov\left(p^*, \frac{dh}{dq} q\right)}{Var(p^*)}} \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq} q\right)}{Var(p^*)}$.

Assumption 4 implies $Cov(\hat{p}, c) = 0$, allowing us to obtain

$$Cov(\hat{p}, \beta(\hat{p} + p^* - c)) = \beta Var(\hat{p})$$
$$Cov(p - p^*, \beta(\hat{p} + p^* - c)) = \beta Var(p) - \beta Var(p^*)$$
$$Var(p)\beta + Cov\left(p, \frac{dh}{dq}q\right) = Var(p^*)\beta + Cov\left(p^*, \frac{dh}{dq}q\right)$$
$$\left(\beta + \frac{Cov\left(p, \frac{dh}{dq}q\right)}{Var(p)}\right)\frac{1}{Var(p^*)} = \left(\beta + \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}\right)\frac{1}{Var(p)}$$
$$\frac{1}{\beta + \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}}\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} = \frac{1}{\beta + \frac{Cov\left(p, \frac{dh}{dq}q\right)}{Var(p)}}\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p)}.$$

Therefore, the probability limit of the OLS estimator can be written as:

$$\text{plim}\hat{\beta}^{OLS} = \beta - \frac{1}{\beta + \frac{Cov\left(p, \frac{dh}{dq}q\right)}{Var(p)}}\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p)}.$$

The roots of the implied quadratic are:

$$\frac{1}{2}\left(\beta^{OLS} - \frac{Cov\left(p, \frac{dh}{dq}q\right)}{Var(p)} \pm \sqrt{\left(\beta^{OLS} + \frac{Cov\left(p, \frac{dh}{dq}q\right)}{Var(p)}\right)^2 + 4\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p)}}\right)$$

which are equivalent to the pair $\left(\beta, \beta\left(1 - \frac{Var(p^*)}{Var(p)}\right) + \frac{Cov(p^*,\xi)}{Var(p^*)} - \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p)}\right)$. Therefore, with the auxiliary condition $\beta < \frac{Cov(p^*,\xi)}{Var(p^*)}\frac{Var(p)}{Var(p^*)} - \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}$, the lower root is consistent for β. QED.

E Supplemental Tables and Figures

Figure E.1 illustrates the connection between rejected values of $Cov(\xi, \eta)$ and rejected values of β that can be obtained from Proposition 5. Panel A of Figure E.1 shows that some intermediate values of $Cov(\xi, \eta)$ can be rejected if $\sigma \geq 0.62$. Uncorrelatedness is rejected with $\sigma \geq 0.69$ and, as $\sigma \to 1$, it must be that $Cov(\xi, \eta) \geq 0.35$. Panel B provides the corresponding bounds on β. As σ becomes larger, a more negative β is required to rationalize the data within the context of the model. If $\sigma = 0.80$ then it must be that $\beta \leq -0.11$. The panel also plots the 3-Stage estimator, $\hat{\beta}^{3-Stage}(\sigma)$, over its supported range. Unlike the bounds, the estimator assumes that $Cov(\xi, \eta) = 0$. As σ converges to 0.69 from below, the 3-Stage estimator approaches the bound on β. For context, the panel plots the upper root of the three-stage estimate, which is negative when $\sigma \geq 0.62$.

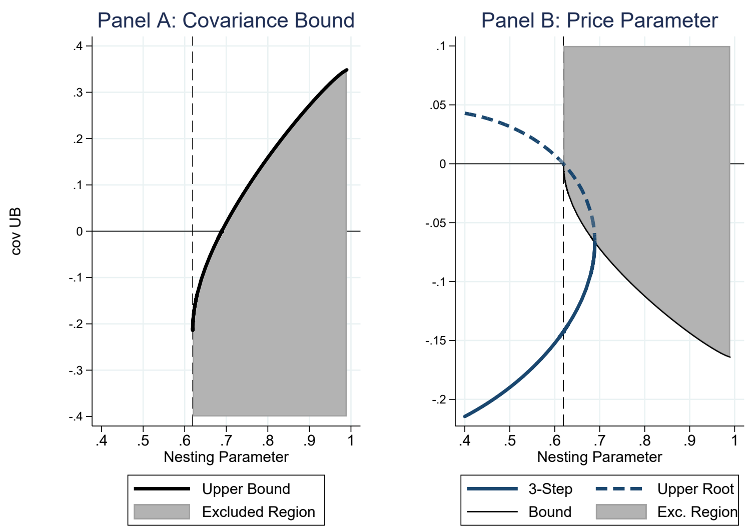

Figure E.1: Prior-Free Bounds for Airlines

Notes: The shaded region in Panel A corresponds to covariance value that cannot be rationalized given the data and the model. The shaded region in Panel B provides the corresponding price parameters that can be ruled out based on the excluded set in Panel A.