Estimating Models of Supply and Demand: Instruments and Covariance Restrictions

Alexander MacKay
Nathan H. Miller

Working Paper 19-051

Estimating Models of Supply and Demand: Instruments and Covariance Restrictions

Alexander MacKay
Harvard Business School

Nathan H. Miller
Georgetown University

Working Paper 19-051

Funding for this research was provided in part by Harvard Business School.

Estimating Models of Supply and Demand: Instruments and Covariance Restrictions*

Alexander MacKay
Harvard University[†]

Nathan H. Miller
Georgetown University[‡]

April 21, 2021

Abstract

We consider the identification of empirical models of supply and demand. As is well known, a supply-side instrument can resolve price endogeneity in demand estimation. We show that, under common assumptions, two other approaches also yield consistent estimates of the joint model: (i) a demand-side instrument, or (ii) a covariance restriction between unobserved demand and cost shocks. The covariance restriction approach can obtain identification even the absence of instruments. Further, supply and demand assumptions alone may bound the structural parameters. We develop an estimator for the covariance restriction approach that is constructed from the output of ordinary least squares regression and performs well in small samples. We illustrate the methodology with applications to ready-to-eat cereal, cement, and airlines.

JEL Codes: C13, C36, D12, D22, D40, L10

Keywords: Identification, Demand Estimation, Covariance Restrictions, Instrumental Variables

*We thank Steven Berry, Charles Murry, Chuck Romeo, Gloria Sheu, Karl Schurter, Jesse Shapiro, Jeff Thurk, Andrew Sweeting, Matthew Weinberg, and Nathan Wilson for helpful comments. We also thank seminar and conference participants at Harvard University, MIT, the University of Maryland, the Barcelona GSE Summer Forum, and the NBER Summer Institute. Previous versions of this paper were circulated with the title "Instrument-Free Demand Estimation."

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

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

1 Introduction

A fundamental challenge in identifying models of supply and demand is that firms can adjust markups in response to demand shocks. Even if marginal costs are constant, this source of price endogeneity generates upward-sloping supply curves in settings with imperfect competition. Thus, the empirical relationship between prices and quantities does not represent a demand curve but rather a mixture of demand and supply. Researchers typically address this challenge by using supply-side instruments to estimate demand, and then employing the formal supply model to recover marginal costs (e.g., Berry et al., 1995; Nevo, 2001).

In this paper, we demonstrate how the supply model can be used to identify models of imperfect competition in the absence of supply-side instruments. We recast equilibrium as a system of simultaneous semi-linear equations that correspond to supply and demand schedules. Through the presence of markups, the endogenous coefficient on price appears in both equations. We use this relationship to formalize two new approaches to identification. First, we show that *demand-side* instruments can also resolve price endogeneity and identify the joint model. Second, we show that a covariance restriction between unobserved demand and cost shocks can fully identify the model without any valid instruments.

Putting our results in context, early research at the Cowles Foundation examined identification in linear systems of equations, including supply and demand models of perfect competition (e.g., Koopmans, 1949; Koopmans et al., 1950).[1] With perfect competition, the supply curve may be upward-sloping due to increasing costs of production, and two separate restrictions are required for identification—one per equation (Hausman and Taylor, 1983). Our contribution lies in the extension to models of imperfect competition, in which markup adjustments affect the slope of supply curves. We demonstrate that demand and markup adjustments are linked theoretically, which reduces the identification requirements. When price endogeneity arises through markups only, we show that a single restriction is sufficient for identification.

We provide formal econometric results for the covariance restrictions approach, which, in the context of imperfect competition, has not previously been examined. In particular, we establish a link between the endogenous price coefficient and the covariance of unobservable cost and demand shocks. Thus, a covariance restriction on unobserved shocks achieves identification. The core intuition is that the supply-side model dictates how prices respond to demand shocks, and thus contains the information necessary to resolve endogeneity bias. With *uncorrelatedness* between demand and cost shocks, the causal price parameter can be recovered from the relative variance of quantity and price in the data. There is no relevance condition that must be satisfied—i.e., no "first-stage" empirical requirement—because the endogenous data

[1]Many articles advanced this research agenda in subsequent decades (e.g., Fisher, 1963, 1965; Wegge, 1965; Rothenberg, 1971; Hausman and Taylor, 1983; Hausman et al., 1987). More recently, Matzkin (2016) examines covariance restrictions in semi-parametric models.

are interpreted directly through the lens of the model.[2]

The results address one of the largest obstacles to applied research in microeconomics: finding valid instruments. Even if such instruments exist in theory, the data to operationalize them may not be available in practice. Our alternative approaches may allow empirical researchers to push ahead along new frontiers. For example, it would be difficult to find the requisite amount of valid instruments for an empirical study of products across many categories. The covariance restriction approach may be appealing in such settings, as it provides a theory-based path to recovering causal parameters without requiring supplemental data.

The strategy of using supply-side restrictions to reduce identification requirements has parallels in a handful of other articles. A simple linear example is provided in Koopmans (1949). Leamer (1981) examines a linear model of perfect competition, and provides conditions under which the price parameters can be bounded using only the endogenous variation in prices and quantities. Feenstra (1994) considers the case of monopolistic competition with constant markups, and a number of application in the trade literature extend this approach (e.g., Broda and Weinstein, 2006, 2010; Soderbery, 2015).[3] Zoutman et al. (2018) return to perfect competition and show that, under a standard assumption in models of taxation, both supply and demand can be estimated with exogenous variation in a single tax rate. Our research builds on these articles by focusing on imperfect competition with adjustable markups.

For much of the paper, we focus on the special case of constant marginal costs, wherein price endogeneity only arises through adjustable markups. This assumption is widespread in empirical models of imperfect competition (e.g., Nevo, 2001; Villas-Boas, 2007; Miller and Weinberg, 2017; Backus et al., 2021). If marginal costs vary with output, then an extra restriction is needed to address simultaneity bias and identify the model. By modeling the relationship between demand and markup adjustments, our results reduce the number of restrictions necessary for identification from three to two.

We organize the paper as follows. Section 2 describes the three approaches to identification, using demand and supply assumptions that are commonly employed in empirical studies of imperfect competition. Section 3 develops an identification strategy using covariance restrictions alone. We show how, even in the absence of supply shifters or demand shifters, a restriction on the correlation of unobservable shocks can achieve point identification. We develop an efficient estimator that can be constructed from the output of ordinary least squares regression. Finally, we compare our covariance restrictions approach to the earlier literature on simultaneous lin-

[2]The international trade literature provides identification results for the special case where markups do not respond to demand shocks (e.g., Feenstra, 1994). We discuss this literature in more detail later. For applications in industrial organization that impose supply-side assumptions, see Thomadsen (2005), Cho et al. (2018), and Li et al. (2021). Thomadsen (2005) assumes no unobserved demand shocks, and Cho et al. (2018) assume no unobserved cost shocks; both are special cases of the covariance restriction approach.

[3]There are interesting historical antecedents to this trade literature. Leamer attributes an early version of his results to Schultz (1928). The identification argument of Feenstra (1994) is also proposed in Leontief (1929). Frisch (1933) provides an important econometric critique.

ear equations. Section 4 provides numerical simulations to compare the three approaches to identification. In small samples, the covariance restriction approach performs well, even when an instrument-based approach suffers from the weak instruments problem. Section 5 provides extensions that examine the robustness of our results under various alternative assumptions about demand and supply. We make use of these extensions in Section 6, which demonstrates the utility of the covariance restrictions approach using three empirical applications.

In Section 7, we discuss the empirical content of supply-side assumptions when combined with the demand model. Our results establish a formal link between the price coefficient, the covariance of unobservable cost and demand shocks, and the empirical variation present in the data. Somewhat surprisingly, we show that supply-side assumptions alone do not improve efficiency when paired with valid cost-shifter instruments. Efficiency gains can be obtained only when supplemental moments—such as covariance restrictions or additional instruments—are added along with the supply-side assumptions. Thus, we provide context for the "folklore" around the Berry et al. (1995) results (Conlon and Gortmaker, 2020). Section 8 concludes.

2 Model

2.1 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, $q_t = [q_{1t}, q_{2t}, \ldots, p_{Jt}]'$, corresponding to each market t, as well as a full rank 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. Let prices and the covariates be linearly independent.

We make the following assumptions about demand and supply:

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

$$h_{jt} \equiv h(q_{jt}, w_{jt}; \sigma) = \beta p_{jt} + x'_{jt}\alpha + \xi_{jt} \tag{1}$$

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

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* (1)*, and has a linear constant marginal cost schedule given by*

$$c_{jt} = x'_{jt}\gamma + \eta_{jt}. \tag{2}$$

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} includes the price and non-price characteristics of other products. For example, with logit demand, $h(q_{jt}; w_{jt}, \sigma) \equiv \ln(s_{jt}/w_{jt})$, where quantities are in shares ($q_{jt} = s_{jt}$), w_{jt} is the share of the outside good (s_{0t}), and there are no additional parameters in σ. Among the other demand systems consistent with Assumption 1 are linear demand, nested logit demand, and random coefficients logit demand; we derive these connections in some detail in Appendix A.

The supply assumption focuses the baseline analysis on Bertrand competition among single-product firms with constant marginal cost. Of these restrictions, only constant marginal cost is consequential for our identification arguments; we maintain the restriction because it is commonly employed in empirical industrial organization. In subsequent sections, we provide the additional notation and results necessary for models with multi-product firms, non-constant marginal costs, and Cournot competition.

We further assume the existence of a Nash equilibrium in pure strategies, such that prices satisfy the first-order condition

$$p_{jt} = c_{jt} - \frac{1}{\beta}\frac{dh_{jt}}{dq_{jt}}q_{jt}. \tag{3}$$

To obtain equation (3), we 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: $p = c - \frac{dp}{dq}q$. With the first order condition in hand, Assumptions 1 and 2 provide a mapping from the data and parameters to the structural error terms (ξ, η). Markups are determined by β, the structure of the model, and observables.[4] As a prelude to our main results, because equilibrium prices respond to demand shocks through markup adjustments, it is possible to model the resulting correlation between prices and demand shocks. This allows for a broader set of identifying restrictions than previously recognized.

In our main analysis, we assume the econometrician seeks to estimate $\theta = (\beta, \alpha, \gamma)$ with knowledge of any nonlinear parameters in σ. Alternatively, the econometrician may consider a candidate σ and wish to obtain corresponding estimates of (β, α, γ), as in the nested fixed-point estimation routine of Berry et al. (1995). Further, as the linear non-price parameters (α, γ) can be recovered trivially given β and σ, we focus our formal identification results on β, the price parameter. In supplementary analyses, we provide an informal discussion of how covariance

[4] We define markups as $p - c$. In particular, markups are proportional to the reciprocal of the price parameter due to the semi-linear demand system. The semi-linear structure is not strictly necessary. In practice, one could start with a known first-order 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}. Our results can be extended to demand systems that admit multiplicative markups (Appendix A).

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.

restrictions can identify σ (Section 5.1) and apply the methodology to estimate the nonlinear parameters of a random coefficient logit model (Section 6.1).

2.2 Three Approaches to Identification

Many empirical applications in industrial organization seek to estimate the demand schedule of equation (1) directly, using instruments taken from the supply-side of the model to address the endogeneity of prices (e.g., Berry et al., 1995; Nevo, 2001; Miller and Weinberg, 2017). A full analysis of the model, however, reveals two addition identification strategies which may be fruitful when supply-side instruments are weak or cannot be found in the data.

By rearranging equation (3), the Nash equilibrium of the oligopoly model can be cast as the solution to two semi-linear simultaneous equations. We provide this result as a lemma.

Lemma 1. ***(Simultaneous Equations Representation)*** *Under assumptions 1 and 2, the equilibrium of the oligopoly model is the solution to the following two simultaneous equations:*

$$\begin{array}{rcll} h_{jt} & = & \beta p_{jt} + x_{jt}'\alpha + \xi_{jt} & \textit{(Demand)} \\ \frac{dh_{jt}}{dq_{jt}} q_{jt} & = & -\beta p_{jt} + x_{jt}'\gamma\beta + \beta\eta_{jt} & \textit{(Supply)} \end{array} \tag{4}$$

where $h_{jt} \equiv h(q_{jt}, w_{jt}; \sigma)$.

In this formulation, the price coefficient appears as a linear coefficient in the right-hand side of the supply equation. Thus, consistent estimation of either demand or supply would suffice to identify the endogenous coefficient, β. With an estimate of β in hand, the other relation is identified and could be estimated with OLS. Because the endogenous variables share a coefficient, it is possible to recover the causal parameters with the addition of a single identifying restriction. We classify different identification strategies as follows:

1. **Supply-side instruments**: A variable that shifts the supply curve but is uncorrelated with ξ may be used as an instrument for price to estimate the demand curve. Price is determined by marginal costs and markups and valid instruments may shift either component. A marginal cost shifter corresponds to a variable $x^{(k)}$ for which $\alpha^{(k)} = 0$ and $\gamma^{(k)} \neq 0$, satisfying the exclusion restriction (in demand) and the relevance condition (in supply). A markup shifter may be obtained from the observable variables in w_{jt}; the characteristics of other products are an example (e.g., Berry et al., 1995).[5] The conditions under which supply-side instruments provide identifying power are well explored in the industrial organization literature (e.g., Berry and Haile, 2014; Reynaert and Verboven, 2014).

[5] Most of the instruments used in empirical industrial organization can be interpreted as supply shifters. For example, prices in other markets can proxy for unobserved marginal costs (e.g., Hausman, 1996; Nevo, 2001), and mergers among competitors can be interpreted as a markup shifter (e.g., Miller and Weinberg, 2017).

2. **Demand-side instruments**: A variable that shifts the demand curve but is uncorrelated with η may be used as an instrument to estimate the estimate supply curve. The inverse representation of the demand system, $p_{jt} = \frac{1}{\beta}h_{jt} - \frac{1}{\beta}x'_{jt}\alpha - \frac{1}{\beta}\xi_{jt}$, suggests two types of instruments. A demand shifter corresponds to a variable $x^{(k)}$ for which $\alpha^{(k)} \neq 0$ and $\gamma^{(k)} = 0$. Alternatively, a markup shifter can be obtained from the observable variables in w_{jt} and used as an instrument. As a supplement to supply-side instruments, Berry et al. (1995) and others have used markup shifters as instruments for the supply equation. To our knowledge, using demand-side instruments as a standalone approach to identification has not yet been recognized in the literature.

3. **Covariance restrictions**: The model is fully identified with a covariance restriction along the lines of $Cov(\xi, \eta) = 0$. The implicit exclusion restrictions are that ξ cannot be in the cost function and η cannot enter demand. Previous articles have demonstrated that, under perfect competition, a covariance restriction must be paired with a demand-side or supply-side instrument to achieve identification (e.g., Hausman and Taylor, 1983; Hausman et al., 1987; Matzkin, 2016). With oligopoly models and constant marginal costs, the covariance restriction alone suffices.

These approaches all impose an orthogonality condition involving one or more unobserved structural error term. There are nonetheless two important distinctions. First, the instrument-based approaches require a relevance condition to be satisfied; by contrast, there is no "first-stage" empirical requirement that must be met with covariance restrictions. Second, covariance restrictions and demand-side instruments require a formal supply-side specification and parametric assumptions on demand, whereas supply-side instruments allow for identification with an informal understanding of supply and nonparametric demand (Berry and Haile, 2014).

Thus, the relative desirability of the three approaches depends on the data availability and the institutional details of the industry under study. We explore the trade-offs in greater detail using small sample Monte Carlo analyses (Section 4). First, however, we demonstrate how a covariance restriction can jointly identify supply and demand, relate our results to the existing literature on simultaneous equations, and discuss the appropriateness of the identifying assumptions.

3 Covariance Restrictions

3.1 Identification

We now formalize the identification argument under a covariance restriction between the unobserved demand and marginal cost shocks. Let β^{OLS} denote the probability limit of the OLS estimate of the price coefficient obtained from a regression of $h(\cdot)$ on p and x. We have

$$\beta^{OLS} \equiv \frac{Cov(p^*, h)}{Var(p^*)} = \beta + \frac{Cov(p^*, \xi)}{Var(p^*)} \tag{5}$$

where $p^* = [I - x(x'x)^{-1}x']p$ is a vector of residuals from a regression of p on x. In the probability limit, the OLS coefficient equals the true parameter plus a bias term.

Two substitutions allow for the bias term to be represented in terms of data, objects that can be estimated with OLS, and the covariance of unobserved demand and cost shocks. First, substitute for price on the right-hand-side of equation (5) using the first-order conditions:

$$\beta^{OLS} = \beta - \frac{1}{\beta}\frac{Cov\left(\frac{dh}{dq}q, \xi\right)}{Var(p^*)} + \frac{Cov(\eta, \xi)}{Var(p^*)}. \tag{6}$$

The second substitution is for ξ in the second term on the right-hand-side, and utilizes that

$$\beta p^*_{jt} + \xi_{jt} = \beta^{OLS} p^*_{jt} + \xi^{OLS}_{jt}$$

where ξ^{OLS} denotes the probability limit of the OLS residuals. Thus, the unobserved demand shock ξ can be expressed in terms of $(p^*, \beta, \beta^{OLS}, \xi^{OLS})$. With this second substitution made, our main identification result obtains after a few additional lines of algebra.

Proposition 1. *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:*

$$\beta^{OLS} = \beta - \frac{1}{\beta + \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}}\frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)} + \beta\frac{1}{\beta + \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}}\frac{Cov(\xi, \eta)}{Var(p^*)}. \tag{7}$$

Therefore, the price parameter β solves the following quadratic equation:

$$\begin{aligned} 0 = \; & \beta^2 \\ & + \left(\frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)} - \beta^{OLS} + \frac{Cov(\xi, \eta)}{Var(p^*)}\right)\beta \\ & + \left(-\beta^{OLS}\frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}\right). \end{aligned} \tag{8}$$

Proof. See appendix.

The proposition establishes a fundamental link between the price coefficient, the covariance of unobservable shocks, and the empirical variation present in the data. With the exceptions of β and $Cov(\xi, \eta)$, all of the terms in equation (8) can be constructed from data. Therefore, the quadratic admits at most two solutions for β for a given value of $Cov(\xi, \eta)$. Thus, assumptions about demand and supply can tightly constrain the set of possible parameter values when combined with prior knowledge of $Cov(\xi, \eta)$. Both possible values can be estimated using the output from OLS and the quadratic formula. As we show next, conditions exist that guarantee point identification and ensure that the lower root of equation (8) is consistent for β.

Proposition 2. ***(Lower Root)*** *Under assumptions 1 and 2, the price parameter β is the lower root of equation* (8) *if and only if the following condition holds:*

$$-\frac{1}{\beta}Cov(\xi, \eta) \leq Cov\left(p^*, -\frac{1}{\beta}\xi\right) + Cov\left(p^*, \eta\right) \tag{9}$$

and, furthermore, β is the lower root of equation (8) *if the following condition holds:*

$$0 \leq \beta^{OLS} Cov\left(p^*, \frac{dh}{dq}q\right) + Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right). \tag{10}$$

Proof. See appendix.

The first (necessary and sufficient) condition depends on the relationship between prices and unobserved shocks. When $Cov(\xi, \eta) = 0$, the condition is violated only if (i) equilibrium prices *decrease* in response to positive demand shocks, and (ii) variation in demand shocks swamps variation in supply shocks. To see this in equation (9), note that the right-hand side depends on the covariance between prices and shocks to inverse demand ($-\frac{1}{\beta}\xi$) as well as the covariance between prices and marginal costs. Because conventional wisdom holds that prices usually increase with positive demand shocks (e.g., Nevo, 2001), we believe it is reasonable to assume β is point identified as the lower root in most empirical applications.[6]

The second (sufficient) condition is derived from the quadratic formula: if the constant term in the quadratic of equation (8) is negative then the upper root of the quadratic is positive and β must be the lower root. Since the terms in (10) are constructed from data, the sufficient condition may be estimated and used to test (and possibly reject) the null hypothesis that multiple negative roots exist.

In the remainder of the our results, we maintain the following assumption:

Assumption 3 (Point Identification): *β is the lower root of equation* (8).

[6]This condition can be verified analytically for some demand systems, including linear demands. Evidence indicating prices increase with observed (positive) demand shifters would also support point identification.

3.2 Estimation with Uncorrelatedness

We develop a three-stage estimator using an orthogonality condition that is standard in the covariance restrictions literature:

Assumption 4 (Uncorrelatedness): $Cov(\xi, \eta) = 0$.

We discuss the applicability of the uncorrelatedness assumption in Section 3.5.

Letting $(\hat{\beta}^{OLS}, \hat{\xi}^{OLS})$ be the OLS estimates of (β, ξ), we have:

Corollary 1. ***(Three-Stage Estimator)*** *Under assumptions 1, 2, 3, and 4, a consistent estimate of the price parameter β is given by*

$$\hat{\beta}^{\text{3-Stage}} = \frac{1}{2}\left(\hat{\beta}^{OLS} - \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)} - \sqrt{\left(\hat{\beta}^{OLS} + \frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\hat{\xi}^{OLS}, \frac{dh}{dq}q\right)}{Var(p^*)}}\right) \tag{11}$$

The estimator is consistent for the lower root of equation (8). 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 is trivial, which may be beneficial if nested inside a search for the σ parameters.[7]

To a first order, the empirical variation that identifies the price parameter is the relative variation in quantity and price in the data. Formally,

Proposition 3. ***(Approximation)*** *A first-order approximation to the three-stage estimator is:*

$$\hat{\beta}^{\text{Approx}} = -\sqrt{\frac{Var(h^*)}{Var(p^*)}} \tag{12}$$

where $h^* = [I - x(x'x)^{-1}x']h$ is a vector of residuals from a regression of h on x.

Intuition for this result can be gleaned from the simultaneous equations representation of the model (Section 2.2), in which β determines the slope of both demand and supply. A large β corresponds to a flatter inverse demand schedule (i.e., price sensitive consumers) and a flatter inverse supply schedule (i.e., less market power). Uncorrelated shifts in such schedules tend to generate more variation in quantity than price. By contrast, a small β corresponds to steeper inverse demand and inverse supply schedules, such that uncorrelated shifts generate more variation in price than quantity. Connecting these observations formally generates an approximation of $\hat{\beta}^{\text{3-Stage}}$ based on the ratio of variances.

[7] A more precise two-stage estimator is available for special cases in which the observed cost and demand shifters are uncorrelated (Appendix B).

An interesting observation, then, is that three-stage estimator effectively converts the endogenous variation in prices and quantity into consistent estimates; it is not necessary to isolate price variation attributable to a valid instrument. This is a fundamental difference between the three-stage estimator and instrument-based estimators (see also Section 3.4), and explains why the three-stage estimator does not require that some relevance condition be satisfied.

The three-stage estimator can be recast as a GMM estimator that exploits the orthogonality condition $E[\xi \cdot \eta] = 0$. Thus, an alternative approach to estimation is to search numerically for a $\tilde{\beta}$ that satisfies the corresponding empirical moment, yielding

$$\hat{\beta}^{GMM} = \arg\min_{\tilde{\beta}<0} \left[\frac{1}{T} \sum_t \frac{1}{|J_t|} \sum_{j \in J_t} \xi(\tilde{\beta}; w_{jt}, \sigma, X_{jt}) \cdot \eta(\tilde{\beta}; w_{jt}, \sigma, X_{jt}) \right]^2 \tag{13}$$

where $\xi(\tilde{\beta}; w, \sigma, X)$ and $\eta(\tilde{\beta}; w, \sigma, X)$ are the estimated residuals given the data and the candidate parameter, and the firms present in each market t are indexed by the set J_t. We see two main situations in which the numerical approach may be preferred despite its greater computational burden. First, additional moments can be incorporated in estimation, allowing for efficiency improvements and specification tests (e.g., Hausman, 1978; Hansen, 1982). Second, the three-stage estimator requires orthogonality between ξ and all regressors (i.e., $E[X\xi] = 0$), whereas the numerical approach can be pursued under a weaker assumption that allows for correlation between ξ and covariates that enter the cost function only.

3.3 Analysis of Bounds

If uncorrelatedness is inappropriate for a particular setting, it may nonetheless be possible to sign $Cov(\xi, \eta)$. This allows for bounds to be place on β. The reason is that there is a one-to-one mapping between the value of $Cov(\xi, \eta)$ and the lower root of equation (8):

Lemma 2. ***(Monotonicity)*** *Under assumptions 1 and 2, a valid lower root of equation* (8) *(i.e., one that is negative) is decreasing in $Cov(\xi, \eta)$. The range of the function is $(0, -\infty)$.*

Proof. See appendix.

Thus, if higher quality products are more expensive to produce ($Cov(\xi, \eta) \geq 0$) or firms invest to lower the marginal costs of their best-selling products ($Cov(\xi, \eta) \leq 0$), then one-sided bounds can be placed on β. More generally, let $r(m)$ be the lower root of the quadratic in equation (8), evaluated at $Cov(\xi, \eta) = m$. Then $Cov(\xi, \eta) \geq m$ produces $\beta \in (-\infty, r(m)]$, and $Cov(\xi, \eta) \leq m$ produces $\beta \in [r(m), 0)$.[8] The lower root, $r(m)$, can be estimated using equation (11).

Interestingly, even if $Cov(\xi, \eta)$ cannot be signed, certain values of β may be possible to rule out because they imply that the data are incompatible with the model. To see why, represent

[8]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.

the quadratic of equation (8) 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 (8), this places restrictions on $Cov(\xi, \eta)$. We now state the result formally:

Proposition 4. ***(Model-Based 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. Further, under assumption 3, this lower bound on $Cov(\xi, \eta)$ provides an upper bound on β.

Proof. See appendix.

From the monotonicity result (Lemma 2), we can use the excluded values of $Cov(\xi, \eta)$ from this result to rule out values of β. Under assumption 3, a model-based upper bound for β is obtained by evaluating the lower root of equation (8) at the model-based bound of $Cov(\xi, \eta)$.

3.4 Relationship to the Simultaneous Equations Literature

To help place these results in context, we provide an overview of the existing literature on the identification of simultaneous equations with covariance restrictions. The subject received early attention in research at the Cowles Foundation (e.g., Koopmans et al., 1950); we focus on the more recent articles of Hausman and Taylor (1983) and Hausman et al. (1987). Adopting their notation, the model of supply and demand is given by:

$$\begin{aligned} y_1 &= \beta_{12}y_2 + \gamma_{11}z_1 + \epsilon_1 && \text{(Supply)} \\ y_2 &= \beta_{21}y_1 + \epsilon_2 && \text{(Demand)} \end{aligned}$$

This system is analogous to equation (4) in the present paper. The most important differences are that (i) by assumption, the covariate z_1 is excluded from the second equation, and (ii) there are two "slope" parameters (β_{12} and β_{21}). The linearity of the system is less consequential; for example, linearity also obtains in our setting with monopoly and linear demands.

Hausman et al. (1987) show that the coefficients β_{12}, β_{21}, and γ_{11} are identified by invoking exogeneity of z_1, the exclusion restriction, and a covariance restriction: $E[z_1 \cdot \epsilon_1] = 0$, $E[z_1 \cdot \epsilon_2] = 0$, and $Cov(\epsilon_1, \epsilon_2) = 0$.[9] The parameters can be estimated jointly via GMM by iterating over candidate parameter values. Alternatively, the demand equation can be estimated

[9] These three equations appear as (2.10a), (2.10b), and (2.10c) in that article.

with 2SLS using z_1 as an instrument, and then the supply equation can be estimated with 2SLS using z_1 and the residual $\hat{\epsilon}_2 = y_2 - \hat{\beta}_{21y_1}$ as instruments (Hausman and Taylor, 1983). Thus, the covariance restriction can be recast as an orthogonality condition involving a residual instrument.[10] Consistent estimation in this context requires that a valid and relevant instrument (z_1) exists, in addition to the covariance restriction on unobserved shocks.[11]

By contrast, the approach that we introduce does not require the presence of an instrument. We use the structure of demand and supply to link the price coefficients β_{12} and β_{21} across the two equations. This reduces the number of endogenous parameters from two to one. Thus, fewer moments are needed for identification, and there is no instrument relevance condition that must be satisfied.

3.5 Assessing Covariance Restrictions

In our experience, economists conducting applied research tend to focus a great deal on the credibility of instruments employed to obtain identification. By contrast, covariance restrictions are more novel, so we provide some thoughts here to help frame future discussions.

We wish to make two observations. The first is that researchers sometimes have detailed knowledge of the determinants of demand and marginal cost, even if these determinants are unobserved in the data. Such knowledge can allow for an assessment of covariance restrictions along the lines of $Cov(\xi, \eta) = 0$. The distinction between *understood* and *observed* is important, as the econometrician may have reasonable priors about the relationship between structural error terms even though they are (by definition) unobserved. To the extent that empirical applications in industrial organization come to rely on covariance restrictions for identification, investments in knowledge about the economic environment become even more important.

Our second observation is that some connections between costs and demand naturally arise in many models. 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. With sufficiently rich data, these sources of confounding variation may be absorbed with control variables or fixed effects. The (residual) structural error terms are orthogonal to these components, so, even with these

[10]This interpretation of covariance restrictions as allowing for residual instruments has been influential. For example, see the lecture notes of Professor Daniel McFadden, dated 1999, which state that:

> Even covariance matrix restrictions can be used in constructing instruments. For example, if you know that the disturbance in an equation you are trying to estimate is uncorrelated with the disturbance in another equation, then you can use a consistently estimated residual from the second equation as an instrument.

The notes can be obtained at `https://eml.berkeley.edu/~mcfadden/e240b_f01/ch6.pdf`, last accessed July 17, 2019.

[11]Matzkin (2016) and Chiappori et al. (2017) provide extensions of this approach to semi-parametric models. Hausman et al. (1987) also consider the identification of simultaneous equations with covariance restrictions alone. This requires at least three equations, however, and thus is not applicable to models of supply and demand.

considerations, $Cov(\xi, \eta) = 0$ can be a valid identifying assumption.

To make this second point explicit, suppose the demand and cost functions are given by:

$$\begin{aligned} h(q_{jt}) &= \beta p_{jt} + x_{jt}'\alpha + D_j + F_t + E_{jt} \\ c_{jt} &= x_{jt}'\gamma + U_j + V_t + W_{jt} \end{aligned}$$

where, again, the subscripts j and t refer to products and markets, respectively. The unobserved shocks are $\xi_{jt} = D_j + F_t + E_{jt}$ and $\eta_{jt} = U_j + V_t + W_{jt}$. 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 $Cov(F_t, V_t) \neq 0$, but market fixed effects can be incorporated to absorb the confounding variation. In this manner, the econometrician may be able to isolate particular components of the unobserved shocks over which a covariance restriction is credible.

4 Small-Sample Performance

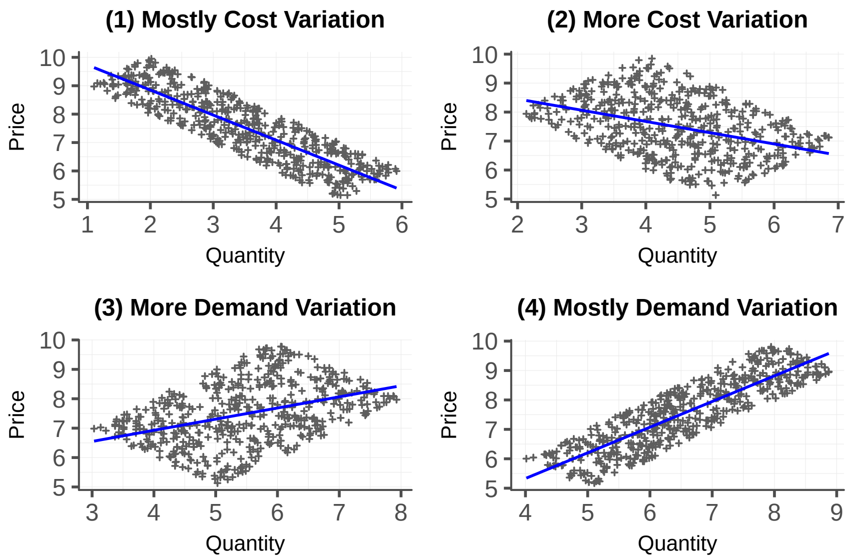
We provide Monte Carlo simulations to illustrate the different approaches to identification and to provide intuition for the covariance restriction approach. First, we consider the impact of relative variation in demand and supply shocks on the performance of the estimators. Second, we consider the impact of the curvature of demand. Finally, we consider misspecification of the supply-side model, when the econometrician makes an assumption about conduct that does not match the data-generating process. We use stylized models for our simulations in all three cases. Covariance restrictions perform reasonably well across the specifications, even when the supply-side conduct is misspecified.

4.1 Relative Variation in Demand and Supply Shocks

Consider the specific version of the model in which a monopolist faces a linear demand schedule, $q_t = 10 - p_t + \xi_t$, and marginal cost is given by $c_t = \eta_t$. Thus, $\beta = -1$. The shocks have independent uniform distributions. We vary the relative support of η and ξ to compare environments in which variation arises more from costs or demand. In specification (1), $\xi \sim U(0, 2)$ and $\eta \sim U(0, 8)$. In specification (2), $\xi \sim U(0, 4)$ and $\eta \sim U(0, 6)$. In specification (3), $\xi \sim U(0, 6)$ and $\eta \sim U(0, 4)$. In specification (4), $\xi \sim U(0, 8)$ and $\eta \sim U(0, 2)$.

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, we present one simulation of 500 observations from each specification in Figure 1, along with the fit of an OLS regression of quantity on price. The probability limits of the OLS

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.

estimate in each scenario are -0.882, -0.385, 0.385, and 0.882. The greater is the demand-side variation, the larger is the endogeneity bias.[12]

We consider the empirical performance of the three-stage estimator in this setting. Regardless of the reduced-form relationship between quantity and price, the covariance restriction approach allows for a consistent estimate of the price parameter by using the relative variation of these outcomes (Proposition 3). In the case of linear demand, the intuition is sharp because the approximation is exact: $plim\left(-\sqrt{\frac{Var(q^*)}{Var(p^*)}}\right) = \beta$. We also estimate demand using two instrument-based approaches: supply shifters and demand shifters. For the supply-shifter approach, we estimate demand with 2SLS using the cost shock η_t as the excluded instrument. That is, we use all of the supply-side variation to estimate demand. For the demand-shifter approach, we estimate the supply equation with 2SLS using ξ_t as the excluded instrument.

To compare the small-sample performance of the approaches, we consider sample sizes of 25, 50, 100, and 500 observations. For each specification and sample size, we randomly draw 10,000 datasets. To account for large coefficients arising from the weak instrument problem, we bound the estimates of β on the range $[-100, 100]$. For specifications that suffer from weak instruments, this will bias the standard errors toward zero. This only affects specifications where the estimated standard error is greater than one, i.e., in 9 of our 48 specifications.

[12]Inspection of Figure 1 further suggests that there may be connection between OLS bias and goodness-of-fit. Indeed, starting with equation (12), a few lines of additional algebra obtain $\beta \approx -\left|\beta^{OLS}\right|/\sqrt{R^2}$ where R^2 is from the residual OLS regression of h^* on p^*. The approximation is exact with linear demand. This reformulation fails if $R^2 = 0$, but numerical results indicate robustness for values of R^2 that are approximately zero. We thank Peter Hull for suggesting this connection.

Table 1: Small-Sample Properties: Relative Variation in Demand and Supply Shocks

(a) Covariance Restrictions

Observations	(1) $Var(\eta) \gg Var(\xi)$		(2) $Var(\eta) > Var(\xi)$		(3) $Var(\eta) < Var(\xi)$		(4) $Var(\eta) \ll Var(\xi)$	
25	-1.004	(0.098)	-1.017	(0.201)	-1.018	(0.206)	-1.005	(0.099)
50	-1.001	(0.068)	-1.008	(0.136)	-1.007	(0.135)	-1.001	(0.068)
100	-1.001	(0.047)	-1.003	(0.094)	-1.004	(0.093)	-1.001	(0.047)
500	-1.000	(0.021)	-1.001	(0.041)	-1.001	(0.042)	-1.000	(0.021)

(b) Supply Shifters (IV-1)

Observations	(1) $Var(\eta) \gg Var(\xi)$		(2) $Var(\eta) > Var(\xi)$		(3) $Var(\eta) < Var(\xi)$		(4) $Var(\eta) \ll Var(\xi)$	
25	-1.004	(0.105)	-1.039	(0.303)	-1.310	(2.629)	-0.899	(13.921)
50	-1.001	(0.072)	-1.018	(0.201)	-1.113	(1.135)	-1.392	(10.890)
100	-1.001	(0.050)	-1.008	(0.138)	-1.048	(0.332)	-1.432	(5.570)
500	-1.000	(0.022)	-1.001	(0.060)	-1.009	(0.138)	-1.061	(0.411)

(c) Demand Shifters (IV-2)

Observations	(1) $Var(\eta) \gg Var(\xi)$		(2) $Var(\eta) > Var(\xi)$		(3) $Var(\eta) < Var(\xi)$		(4) $Var(\eta) \ll Var(\xi)$	
25	-0.881	(12.794)	-1.295	(3.087)	-1.040	(0.312)	-1.006	(0.106)
50	-1.448	(10.980)	-1.112	(0.596)	-1.016	(0.198)	-1.001	(0.073)
100	-1.597	(5.837)	-1.045	(0.333)	-1.009	(0.136)	-1.001	(0.050)
500	-1.070	(0.414)	-1.008	(0.137)	-1.002	(0.060)	-1.000	(0.022)

Notes: Results are based on 10,000 simulations of data for each specification and number of observations. The demand curve is $q_t = 10 - p_t + \xi_t$, so that $\beta = -1$, and marginal costs are $c_t = \eta_t$. IV-1 estimates are calculated using two-stage least squares with marginal costs (η) as an instrument in the demand equation. Analogously, IV-2 estimates are calculated using two-stage least squares with demand shocks (ξ) as an instrument in the supply equation. In specification (1), $\xi \sim U(0,2)$ and $\eta \sim U(0,8)$. In specification (2), $\xi \sim U(0,4)$ and $\eta \sim U(0,6)$. In specification (3), $\xi \sim U(0,6)$ and $\eta \sim U(0,4)$. In specification (4), $\xi \sim U(0,8)$ and $\eta \sim U(0,2)$.

Table 1 summarizes the results. Panel (a) considers the covariance restriction approach. Because the estimator exploits all of the variation in the data, its performance does not depend on whether empirical variation arises more from variation in cost conditions or demand conditions. This can be seen by comparing specifications (1) and (2) (primarily cost variation) to specifications (3) and (4) (primarily demand variation). In particular, the standard errors are symmetric for specifications (1) and (4), as well as for specifications (2) and (3). Comparing estimates across panels, we see that the three-stage estimator is more efficient and has less bias than the instrument-based estimators, across all specifications and sample sizes considered.

Panel (b) considers the supply-shifter approach. When variation in costs dominate variation in demand shocks (as in the leftmost column), cost-based instruments perform well. Performance deteriorates as relatively more variation is caused by demand shocks, and this is exacerbated in smaller samples. The large bias in the upper right of panel (b) reflects a weak instrument problem, even though all the exogenous cost variation in observed. With specification (4), the mean F-statistics for the first-stage regression of p on η are 2.6, 4.2, 7.3, and 32.6 for markets with 25, 50, 100, and 500 observations, respectively. As this approach exploits only the price variation that can be attributed to costs, the success of the approach depends on the relative importance of cost to demand shocks in the data-generating process.[13]

Panel (c) considers the demand-shifter approach. The rightmost column shows that the demand-shifter approach performs well when the variation in demand shocks dominates the variation in cost shocks. The performance of this approach deteriorates as relatively more variation in the model is attributable to costs. Where the supply-shifter approach does very well, the demand-shifter approach does poorly, and vice versa.

All three approaches rely on the orthogonality condition $Cov(\xi, \eta) = 0$, which is equivalent to the the exclusion restriction for the instrument-based approaches. Though the assumptions about the relations among unobservables overlap in this example, the instrument-based approaches also require the econometrician to observe either η_t or ξ_t. The covariance restriction approach is consistent when both are unobserved. Therefore, in settings where instrumental variables perform poorly, a three-stage estimator may still provide a precise estimate.

4.2 Curvature of Demand

We next consider how the perfomance of the estimation approaches depends on the curvature of demand. We extend the stylized model above to allow the demand function to take a non-linear form: $h(q_t; \sigma) = 10 - p_t + \xi_t$. We parameterize the semi-linear demand schedule as $h(q; \sigma) = \frac{1}{\sigma} q^{\sigma}$. When $\sigma = 1$, we obtain the linear specification from the previous set of simulations. When $\sigma > 1$, the demand curve is concave, and when $\sigma < 1$, it is convex. We consider five specifications: $\sigma \in \{\frac{1}{3}, \frac{1}{2}, 1, 2, 3\}$. In each, we assume ξ and η are have equal variance ($\sim U(0, 5)$).

Table 2 reports mean coefficients and standard errors for the three estimation approaches. The covariance restriction approach is robust to changes in the curvature of demand, as the mean coefficients and standard errors are essentially unchanged across specifications. Conversely, the performance of the instrument-based estimators depends on the demand curvature. Supply shifters (IV-1) perform best with convex demand, as the first column has the smallest standard errors and least bias. As demand becomes more concave, the supply-side instruments become less reliable. In specifications (4) and (5), supply shifters generate biased estimates and substantially larger standard errors. Analogously, the demand shifter approach (IV-2) performs best with concave demand and degrades as demand becomes more convex.

[13] Additionally, specifications (3) and (4) generate positive OLS coefficients. In small samples, a positive OLS coefficient might suggest that an identification strategy relying on cost shifters or other supply-side instruments will not be fruitful, because much of the variation in prices is determined by demand. Of course, for any particular instrument, the first-stage F-statistic can be used to assess relevance.

Table 2: Small-Sample Properties: Demand Curvature

Estimation Method	(1) $\sigma = 1/3$	(2) $\sigma = 1/2$	(3) $\sigma = 1$	(4) $\sigma = 2$	(5) $\sigma = 3$
Covariance Restrictions	-1.002	-1.003	-1.003	-1.003	-1.002
	(0.072)	(0.072)	(0.072)	(0.072)	(0.072)
IV-1: Supply Shifters	-1.003	-1.005	-1.011	-1.034	-1.076
	(0.096)	(0.109)	(0.147)	(0.237)	(0.386)
IV-2: Demand Shifters	-1.076	-1.034	-1.011	-1.004	-1.003
	(0.393)	(0.237)	(0.147)	(0.108)	(0.096)
IV-1: First-stage F-statistic	1797.5	799.4	200.6	50.9	23.2
IV-2: First-stage F-statistic	23.2	51.0	201.0	801.0	1801.3
Cost Pass-through	0.750	0.667	0.500	0.333	0.250

Notes: Results are based on 10,000 simulations of 200 observations for each specification. The demand curve is $h_t = 10 - p_t + \xi_t$, so that $\beta = -1$, and marginal costs are $c_t = \eta_t$. IV-1 estimates are calculated using two-stage least squares with marginal costs (η) as an instrument in the demand equation. Analogously, IV-2 estimates are calculated using two-stage least squares with demand shocks (ξ) as an instrument in the supply equation. Across all specifications, $\xi \sim U(0,5)$ and $\eta \sim U(0,5)$. The curvature of demand varies across specifications according to the parameter σ: $h(q;\sigma) = \frac{1}{\sigma}q^{\sigma}$.

These results parallel the previous simulations which vary the relative contribution of supply and demand shocks. Though the magnitudes of these shocks are indexed to prices,[14] the curvature of demand dictates how these shocks get transformed into quantities, and, through the equilibrium response, how these shocks pass-through to prices.

To highlight this, we provide the mean cost pass-through in the final row of Table 2. Cost pass-through is the highest when demand is convex (0.750 when $\sigma = \frac{1}{3}$) and falls with more concave demand. Thus, the same marginal cost shock has smaller impact on equilibrium prices when demand is more concave. This has implications for the relevance of the instrument. The first-stage F-statistics for supply-side instruments, which are also reported in the table, decline with more concave demand. Intuitively, supply-side instruments are weaker as when cost pass-through is lower. Symmetric logic applies to demand-side instruments.

4.3 Supply-Side Misspecification

To illustrate how supply-side misspecification may affect the performance of the estimators, we simulate duopoly markets in which the standard assumption of Bertrand price competition may not match the data-generating process. We assume the demand system is logit, providing consumers with a differentiated discrete choice, and we allow them to select an outside option

[14] Supply shocks (η) have the same coefficient as price in the supply equation, and demand shocks (ξ) have the opposite sign, but same coefficient magnitude, as price in the demand equation: $h(q_t;\sigma) = 10 - p_t + \xi_t$. The demand equation relationship holds because $\beta = -1$.

Table 3: Small-Sample Properties: Supply-Side Misspecification

Estimation Method	(1) $\kappa = 0.0$	(2) $\kappa = 0.2$	(3) $\kappa = 0.4$	(4) $\kappa = 0.6$	(5) $\kappa = 0.8$	(6) $\kappa = 1.0$
Covariance Restrictions	-1.001	-1.002	-1.000	-1.003	-1.016	-1.038
	(0.050)	(0.052)	(0.053)	(0.054)	(0.053)	(0.051)
IV-1: Supply Shifters	-1.002	-1.000	-1.001	-1.001	-1.001	-1.002
	(0.076)	(0.077)	(0.077)	(0.076)	(0.073)	(0.071)
IV-2: Demand Shifters	-1.015	-1.017	-1.012	-1.025	-1.082	-1.220
	(0.153)	(0.155)	(0.159)	(0.178)	(0.213)	(0.298)
IV-1: First-stage *F*-statistic	1079.9	1335.3	1424.9	1277.8	1027.1	801.9
IV-2: First-stage *F*-statistic	99.0	108.4	111.4	100.8	77.8	50.8

Notes: Results are based on 10,000 simulations of 200 duopoly markets for each specification. The demand curve is $h_{jt} = 2 - p_{jt} + \xi_{jt}$, so that $\beta = -1$, and marginal costs are $c_{jt} = \eta_{jt}$. Demand is logit: $h(q_{jt}) = \ln(q_{jt}) - \ln(q_{0t})$, where q_{0t} is consumption of the outside good. IV-1 estimates are calculated using two-stage least squares with marginal costs (η) as an instrument in the demand equation. Analogously, IV-2 estimates are calculated using two-stage least squares with demand shocks (ξ) as an instrument in the supply equation. Across all specifications, $\xi \sim U(0, 0.5)$ and $\eta \sim U(0, 0.5)$. The data-generating process varies in the nature of competition across specifications, indexed by the conduct parameter κ. The coefficients are estimated under the (misspecified) assumption of Bertrand price competition ($\kappa = 0$).

in addition to a product from each firm. The quantity demanded of firm j in market t is

$$q_{jt} = \frac{\exp(2 - p_{jt} + \xi_{jt})}{1 + \sum_{k=j,i} \exp(2 - p_{kt} + \xi_{kt})}$$

On the supply-side, marginal costs are $c_{kt} = \eta_{kt}$ ($k = j, i$). Firm j sets price to maximize $\pi_j + \kappa \pi_i$, and likewise for firm i, where $\kappa \in [0, 1]$ is a conduct parameter (e.g., Miller and Weinberg, 2017). The first-order conditions take the form

$$\begin{bmatrix} p_j \\ p_i \end{bmatrix} = \begin{bmatrix} c_j \\ c_i \end{bmatrix} - \left[\begin{pmatrix} 1 & \kappa \\ \kappa & 1 \end{pmatrix} \circ \left(\frac{\partial q}{\partial p} \right)^T \right]^{-1} \begin{bmatrix} q_j \\ q_i \end{bmatrix}$$

where $\frac{\partial q}{\partial p}$ is a matrix of demand derivatives and $\circ$ denotes element-by-element multiplication. The model nests Bertrand competition ($\kappa = 0$) and joint price-setting behavior ($\kappa = 1$), as well as capturing (non-micro-founded) intermediate cases by scaling markups. Of course, there are other ways to change the equilibrium concept; these simulations are not fully general.

We generate data with different conduct parameters: $\kappa \in \{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$. For each specification, we simulate datasets with 400 observations (200 markets $\times$ two firms), and estimate the model under the erroneous assumption of Bertrand price competition ($\kappa = 0$), thus generating supply-side misspecification. Table 3 displays the results.

As expected, supply-side misspecification introduces bias into the covariance restrictions approach, with a bias of -3.8 percent when the true nature of conduct is $\kappa = 1$. The bias does not appear to be meaningful for modest values of κ (i.e., 0.4 or less). Likewise, the demand-side instruments (IV-2), which invoke the formal assumption about conduct in estimation, perform worse when the true κ is farther from the assumed value. The demand-side instruments perform poorly when the true conduct is $\kappa = 1$, with a mean bias of over 20 percent.

By contrast, supply-side instruments do not use a formal assumption about conduct in estimation and provide consistent estimates across the specifications (IV-1). Consistent with the earlier simulations, the three-stage estimator outperforms IV-1 when conduct is correctly specified ($\kappa = 0$). These results illustrate a key trade-off to the econometrician: if the supply-side assumptions are to be maintained, then covariance restrictions can offer better precision relative to instrument-based approaches. However, supply-side instruments are robust to misspecification of firm conduct, whereas covariance restrictions are not.

We note that the covariance restriction approach, which uses both demand-side and supply-side variation, is not as susceptible to misspecification bias as demand-side instruments in our simulations. The estimator appears to place greater weight on the source of variation with more power. In specification (6), the mean coefficient of -1.038 is much closer to the supply-shifter mean of -1.002 than the demand-shifter mean of -1.220. Indeed, it is approximately equal to the IV-1 and IV-2 estimates weighted by the square root of the respective F-statistics. By placing greater weight on supply-side shocks as the demand-side instruments degrade, the covariance restriction approach receives some protection against bias from model misspecification.

5 Extensions and Generalizations

The results developed thus far rely on some relatively strong (though common) restrictions on the form of demand and supply. In this section, we consider generalizations to non-constant marginal costs and alternative models of competition. First, we provide guidance on how additional restrictions may be imposed to estimate demand systems when σ is unknown. We provide additional extensions the appendices. In Appendix A, we discuss how to generalize the demand assumption to incorporate, for example, constant elasticity demand. In Appendix C, we provide the extension to multi-product firms.

5.1 Identification of Nonlinear Parameters

We now consider identification of the parameter vector, σ, which contains parameters that enter the demand system nonlinearly or load onto endogenous regressors other than price. Conditional on σ, knowledge of $Cov(\xi, \eta)$ can be sufficient to identify the linear parameters in the model, including β (Proposition 2). Thus, the single covariance restriction provides a function that maps σ to β and generates an identified set for (β, σ). To point identify (β, σ) when

σ is unknown, supplemental moments must be employed. These moments may be generated from instruments or from covariance restrictions that extend the notion of uncorrelatedness.

Consider the relation between any demand shock and any marginal cost shock across products (j, k) and markets (t, s): (ξ_{jt}, η_{ks}). The slightly stronger form of the uncorrelatedness assumption holds within the observations of each product:

$$E_t[\xi_{jt} \cdot \eta_{jt}] = 0 \quad \forall j, \tag{14}$$

where the expectation is taken over markets. If this stronger form of uncorrelatedness is credible then the additional restrictions

$$E_t[\xi_{jt} \cdot \eta_{kt}] = 0 \quad \forall j, k \tag{15}$$

also are likely to be credible. Because the expectation in equation (15) is taken over markets for each pair (j, k), there are $J \times J$ potentially identifying restrictions. In specific applications, with more structure on the model, it may be possible to verify point identification. As the moment conditions we propose are nonlinear in σ, in the general case we do not rule out that the corresponding loss function could feature multiple local minima.

We impose this set of moments in our first application, in which we estimate 12 nonlinear parameters in the random coefficients demand system of Nevo (2000). Roughly, cross-product covariance restrictions allow the empirical relationship between the quantity of each product j and the prices of each product k to be interpreted as arising from model parameters (σ), rather than from systematic correlation between the demand shocks of one product and the cost shocks of another. However, cross-covariance restrictions between products represent only one possible path to identifying the (β,σ) pair; many other restrictions are available.[15]

5.2 Non-Constant Marginal Costs

Thus far, we have examined marginal costs that are constant in output, an assumption that is maintained in many empirical applications (e.g., Nevo, 2001; Miller and Weinberg, 2017). This restriction is necessary if the demand and supply schedules, obtained in the simultaneous equations representation of the model, are to have slopes of identical magnitude. With non-constant marginal costs, identification requires an additional moment restriction.

[15] Two examples suffice. First, it may be reasonable to assume that the variance of the demand shock does not depend on the level of the cost shock, and vice versa, which generates the moments $E_{jt}[\xi_{jt}^2 \cdot \eta_{jt}]$ and $E_{jt}[\xi_{jt} \cdot \eta_{jt}^2]$. Second, it may be reasonable to assume that shocks are uncorrelated when aggregated by product group:

$$E_{gt}[\overline{\xi}_{gt} \cdot \overline{\eta}_{gt}] = 0,$$

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.

To make this observation explicit, let marginal costs take the form:

$$c_{jt} = x'_{jt}\gamma + g(q_{jt};\lambda) + \eta_{jt} \tag{16}$$

Here, $g(q_{jt};\lambda)$ is a potentially nonlinear function. 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, markup adjustments are no longer the sole mechanism through with prices respond to demand shocks. The OLS regression of $h(q_{jt}, w_{jt};\sigma)$ on p and x yields a price coefficient with the following probability limit:

$$plim(\hat{\beta}^{OLS}) = \beta - \frac{1}{\beta}\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)} + \frac{Cov(\xi, g(q;\lambda))}{Var(p^*)} \tag{17}$$

The third term on the right-hand-side shows that bias depends on how demand shocks affect the non-constant potion of marginal costs. Without prior knowledge of $g(q_{jt};\lambda)$, an additional restriction is necessary to extend the identification results of the preceding sections. For any candidate $\hat{\lambda}$, however, a corresponding three-stage estimator of $\beta(\hat{\lambda})$ can be obtained:

Proposition 5. *Under assumptions 1 and a modified assumption 2 in which marginal costs take the semi-linear form of equation* (16)*, β solves the following quadratic equation:*

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

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

Proof. See appendix.

To identify the cost function, additional covariance restrictions—such as those discussed in the preceding section—or other instruments may be employed to pin down λ.

5.3 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. We provide one such extension in the empirical application to the cement industry.

The approach can be extended to demand systems that generate an alternative structure for equilibrium prices. For example, consider a monopolist facing a constant elasticity demand schedule. The optimal price is $p = c\frac{\epsilon}{1+\epsilon}$ where $\epsilon < 0$ is the elasticity of demand. In this model, the monopolist has a multiplicative markup that does not respond to demand shocks. As we show in Appendix A, it is straightforward to extend our identification results to this model and other demand systems that admit a more general class of multiplicative markups, which result when demand is semi-linear in log prices.[16]

6 Empirical Applications

6.1 Ready-to-Eat Cereal

In our first application, we examine the pseudo-real cereals data of Nevo (2000).[17] The model features random coefficients logit demand and Bertrand competition among multi-product firms. We use the application to demonstrate that covariance restrictions can identify the price parameter, β, as well as the nonlinear parameters, σ. In the data, there is no variation in the

[16] Our identification result relies on separability between marginal costs and markups, which may be obtained after a transformation of prices.

[17] See also Dubé et al. (2012), Knittel and Metaxoglou (2014), and Conlon and Gortmaker (2020). We focus on the "restricted" specification of Conlon and Gortmaker (2020), which addresses a multicollinearity problem by imposing that the nonlinear parameter on Price×Income^2 takes a value of zero.

product choice sets across markets, so the instruments of Berry et al. (1995) are unavailable. The instruments provided in the data set and employed in Nevo (2000) are constructed from the prices of the same product in other markets. We compare the estimates from a specification with these instruments to a specification with covariance restrictions.

The indirect utility that consumer i receives from product j in market m and period t is given by

$$u_{ijmt} \quad = \quad x_j \alpha_i^* + \beta_i^* p_{jmt} + \zeta_j + \xi_{jmt} + \epsilon_{ijmt}$$

where ζ_j is a product fixed effect and ϵ_{ijmt} is a logit error term. The indirect utility provided by the outside good, $j = 0$, is $u_{i0mt} = \epsilon_{i0mt}$. The consumer-specific coefficients take the form

$$\begin{bmatrix} \alpha_i^* \\ \beta_i^* \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \Pi D_i + \Sigma \nu_i$$

where D_i is a vector of observed demographics and ν_i is vector of unobserved demographics that have independent standard normal distributions. Within the notation of Section 2, we have $\sigma = (\Pi, \Sigma)$ and $w_{jmt} = (D, \nu)$. The transformation of quantities, $h(q_{jmt}, w_{jmt}; \tilde{\sigma})$, can be recovered using the contraction mapping of Berry et al. (1995) for any $\tilde{\sigma}$. The supply side of the model is the multiproduct version of our Assumption 2 (see Appendix C).

The data are a panel of 24 brands, 47 markets, and two quarters. We estimate the demand parameters, $\theta = (\beta, \alpha, \Pi, \Sigma)$, using the covariance restrictions $Cov(\xi_j, \eta_k) = 0$ for all j, k, as proposed in Section 5.1. The restrictions are valid if the demand shock of each product is orthogonal to its own marginal cost shock and those of all other products. The $J \times J$ (= 576) covariance restrictions are sufficient for estimation. We assume that point identification of the twelve parameters in σ is achieved.[18]

Table 4 summarizes the results of estimation based on the instruments (panel (a)) and covariance restrictions (panel (b)). Both identification strategies yield similar mean own-price demand elasticities: -3.70 with instruments and -3.61 with covariance restrictions. Overall, the different approaches produce similar patterns for the coefficients. Most of the point estimates under covariance restrictions fall in the 95 percent confidence intervals implied by the specification with instruments, including that of the mean price parameter. Only one of the interaction terms that is statistically significant (Constant×Income) changes sign. Excluding the standard deviation parameters, the standard errors are noticeably smaller with covariance restrictions, which likely reflects the greater number of identifying restrictions. Finally, covariance restrictions fit the data reasonably well. At the estimated parameters, the correlation between own demand shocks and own marginal cost shocks is 0.0015 and the mean absolute

[18] In many applications of the random coefficients demand model, informal checks for identification are conducted by estimating with many different starting values (e.g., Miller and Weinberg, 2017). Those exercises remain valuable if instruments are replaced with covariance restrictions.

Table 4: Point Estimates for Ready-to-Eat Cereal

(a) Available Instruments

Variable	Means	Standard Deviations	Interactions with Demographics: Income	Age	Child
Price	-32.019	1.803	4.187	–	11.755
	(2.304)	(0.920)	(4.638)		(5.198)
Constant	–	0.120	3.101	1.198	–
		(0.163)	(1.105)	(1.048)	
Sugar	–	0.004	-0.190	0.028	–
		(0.012)	(0.035)	(0.032)	
Mushy	–	0.086	1.495	-1.539	–
		(0.193)	(0.648)	(1.107)	

(b) Covariance Restrictions

Variable	Means	Standard Deviations	Interactions with Demographics: Income	Age	Child
Price	-36.230	1.098	14.345	–	26.906
	(1.122)	(1.067)	(1.677)		(1.384)
Constant	–	0.051	-0.156	1.072	–
		(0.230)	(0.286)	(0.240)	
Sugar	–	0.003	-0.084	-0.004	–
		(0.014)	(0.018)	(0.010)	
Mushy	–	0.130	0.301	-0.845	–
		(0.162)	(0.196)	(0.103)	

Notes: Table reports point estimates for the random-coefficients logit demand system estimated using the Nevo (2000) dataset. Panel A employs the available instruments and replicates Conlon and Gortmaker (2020). Panel B employs covariance restrictions described by the text.

value across the 576 sample moments is 0.1145.

6.2 The Portland Cement Industry

Our second 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.[19] We use the application to demonstrate that knowledge of institutional details can help evaluate the uncorrelatedness assumption.

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

[19] 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 (16), 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:

$$plim\left(\hat{\beta}^{OLS}\right) = \beta - \frac{1}{\beta}\frac{1}{J}\frac{Cov\left(\xi, \frac{dh}{dq}Q\right)}{Var(p^*)} + \frac{Cov(\xi,\overline{g})}{Var(p^*)} \tag{18}$$

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 $g(q;\lambda)$.[20] The identification result provided in Section 5.2 for models with non-constant marginal costs extends.

Corollary 2. *In the Cournot model, the price parameter β solves the following quadratic equation:*

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

The derivation tracks exactly the proof of Proposition 5. 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 markets in the United States annually over 1984-2009. Let the demand curve in market m and year t have a logit form:

$$h(Q_{mt};w) \equiv \ln(Q_{mt}) - \ln(M_m - Q_{mt}) = \alpha_r + \beta p_{mt} + \xi_{mt} \tag{19}$$

where M_m is the "market size" of the market. We assume $M_r = 2 \times \max_t\{Q_{mt}\}$ for simplicity.[21] Further, let marginal costs take the "hockey stick" form of FRR:

$$\begin{aligned}
c_{jmt} &= \gamma + g(q_{jmt}) + \eta_{jmt} \\
g(q_{jmt}) &= 2\lambda_2 1\{q_{jmt}/k_{jm} > \lambda_1\}(q_{jmt}/k_{jm} - \lambda_1)
\end{aligned} \tag{20}$$

[20] 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.

[21] 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.

Table 5: 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 market-year observations over 1984-2009. Bootstrapped standard errors are based on 200 random samples constructed by drawing markets with replacement.

where k_{jm} and q_{jmt}/k_{jm} are capacity and utilization, respectively. Marginal costs are constant 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. Indeed, this is precisely the identification argument of FRR, as both coal and electricity prices are included in the set of excluded instruments.

Table 5 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 with additional restrictions (see Section 5.2), but the 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 (2SLS) or unobserved (3-Stage).

6.3 The Airline Industry

In our third empirical application, we examine demand for airline travel using the setting and data of Aguirregabiria and Ho (2012) ["AH"].[22] 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. We use the application to demonstrate how bounds can narrow the identified set for (β, σ).

The nested logit demand system can be expressed as

$$h(s_{jmt}, w_{jmt}; \sigma) \equiv \ln s_{jmt} - \ln s_{0mt} - \sigma \ln \overline{s}_{jmt|g} = \beta p_{jmt} + x'_{jmt}\alpha + \xi_{jmt} \tag{21}$$

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.

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. Markets are defined as directional round trips between origin and destination cities. 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. The supply side of the model is the multiproduct version of Assumption 2 (see Appendix C).[23]

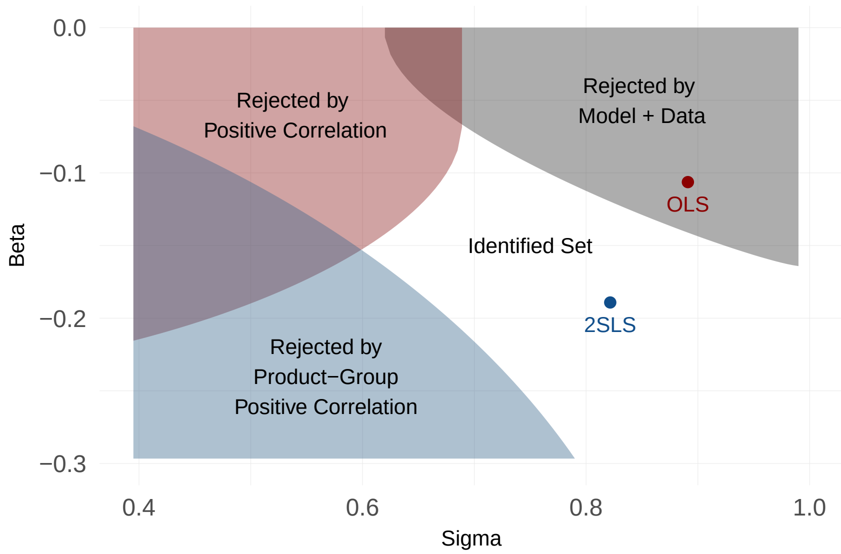
The institutional details of the airline industry suggest that the covariance between demand shocks and marginal cost shocks is positive. The marginal production cost to carry an additional passenger is small and roughly constant because each additional passenger has little impact on the inputs needed to fly the plane from one airport to another. However, the airline bears an opportunity cost for each sold seat, as they can no longer sell the seat at a higher price to another passenger. When the airline expects the flight to be at capacity, this opportunity cost may become large (Williams, 2020). Thus, positive shocks to demand produce more full flights, which motivates the inequality assumption $Cov(\xi, \eta) \geq 0$.

Under that assumption, we reject values (β, σ) that produce a negative correlation in product-specific shocks. We combine this with model-based bounds (Section 3.3). Finally, if the corre-

[22] 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.

[23] 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×quarter fixed effects. The latter expands on the city×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.

Figure 2: Analysis of Bounds in the Airlines Industry



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 $Cov(\overline{\xi}, \overline{\eta}) < 0$. 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. The parameter σ can only take values on $[0, 1)$.

lation in product-level shocks is weakly positive, it is reasonable to assume that the group-level shocks are also weakly positive, through a similar deduction. Thus, we apply the group-level inequality

$$E_{gmt}[\overline{\xi}_{gmt} \cdot \overline{\eta}_{gmt}] \geq 0, \tag{22}$$

where $\overline{\xi}_{gmt} = \frac{1}{|g|} \sum_{j \in g} \xi_{jmt}$ and $\overline{\eta}_{gmt} = \frac{1}{|g|} \sum_{j \in g} \eta_{jmt}$ are the mean demand and cost shocks within a group-market-period. By rejecting parameter values that fail to generate the data or that deliver negative correlations between costs and demand, we narrow the identified set.

Figure 2 displays the rejected regions based on the model and our assumptions on unobserved shocks. The gray region corresponds to the parameter values rejected by the model-based bounds; the model itself rejects some values of β if $\sigma \geq 0.62$. As σ becomes larger, a more negative β is required to rationalize the data. The dark red region corresponds to parameter values that generate negative correlation between demand and supply shocks. The region is rejected under the prior that $Cov(\xi, \eta) \geq 0$. The dark blue region provides the corresponding set for the prior $Cov(\overline{\xi}, \overline{\eta}) \geq 0$ and is similarly rejected.

The three regions overlap, but no region is a subset of another. The non-rejected values provide the identified set. We 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 2. The OLS estimate falls in a rejected region and can be ruled out by the model alone. The 2SLS estimate falls within the identified set. This result is not mechanical, as these point estimates are generated with non-nested assumptions.

7 The Empirical Content of Supply-Side Restrictions

Our analysis shows that imposing a formal supply-side model in estimation expands the range of feasible identifying assumptions. Demand-side instruments or covariance restrictions can be used as standalone strategies to resolve price endogeneity. An outstanding question is whether the supply model has any benefit in estimation without such moments. Our analysis of the joint model provides the answer: supply-side assumptions typically provide *no additional information* if demand-side instruments or covariance restrictions are not employed.

To understand why supply restrictions alone do not generally assist with identification, observe that equation (8) provides a function linking the endogenous parameter β to $Cov(\xi, \eta)$. Inverting this function obtains the implied value of $Cov(\xi, \eta)$ for any candidate value of β:

Corollary 3. ***(Implied Covariance)*** *There is a function mapping β to $Cov(\xi, \eta)$, given by*

$$\begin{aligned} Cov(\xi,\eta) = {} & Var(p^*)\beta^{OLS} - Cov\left(p^*, \frac{dh}{dq}q\right) - \beta Var(p^*) \\ & + \frac{1}{\beta}\left(\beta^{OLS} Cov\left(p^*, \frac{dh}{dq}q\right) + Cov\left(\xi^{OLS}, \frac{dh}{dq}q\right)\right). \end{aligned} \tag{23}$$

Thus, $Cov(\xi, \eta)$ acts as a free parameter that rationalizes an estimate of β, conditional on the data and the structure of demand and supply. The only exception arises in the special case of covariance bounds, where the supply-side can rule out some values of β (Proposition 4).

Efficiency gains are possible if additional moments are imposed along with the supply-side assumptions. One prominent example is Berry et al. (1995), which estimates both the demand and supply relations from equation (4) with markup shifters as instrumental variables, combining the first two approaches enumerated in Section 2.2. The efficiency gains from joint estimation can be attributed to the additional moment conditions. If marginal cost shifters are the only available instruments (e.g., Nevo, 2001) then imposing the supply-side does not yield efficiency improvements because cost shifters do not separately identify the supply equation.

Efficiency gains are also possible through restrictions on the covariance structure of unobservables. Some applications that use the "optimal instruments" method make the assumption of homoskedasticity in the unobserved demand and cost shocks, which can provide modest efficiency gains (Reynaert and Verboven, 2014; Conlon and Gortmaker, 2020).

8 Conclusion

Our objective has been to evaluate the identifying power of supply-side assumptions in models of imperfect competition. Invoking the supply model in estimation expands the set of restrictions that obtain identification. In particular, many applications in industrial organization currently employ instruments from the supply-side of the model in estimation. We show that demand-side instruments and covariance restrictions between unobserved demand and cost shocks also can resolve price endogeneity and allow for consistent estimation. The covariance restrictions approach is notable in part because there is no relevance condition; instead the endogenous variation in quantity and price is interpreted through the lens of the model to recover the structural parameters. As this is somewhat novel, we provide three empirical applications to demonstrate how covariance restrictions can be applied and evaluated.

We view the relative desirability of supply-side instruments, demand-side instruments, and covariance restrictions as depending primarily on data availability and the institutional details of the industry under study. The main advantage of supply-side instruments is that only an informal understanding of supply is required in demand estimation, which allows for robustness in the event of supply-side misspecification. By contrast, demand-side instruments and covariance restriction require a formal supply-side model. Nonetheless, these approaches provide paths to identification that may facilitate research in areas for which strong supply-side instruments are unavailable. The reliability of research that employs these strategies—as is true with most empirical work—depends on the appropriateness of the identifying assumptions.

References

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

Backus, M., C. Conlon, and M. Sinkinson (2021). Common ownership and competition in the ready-to-eat cereal industry. *Working Paper*.

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). 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.

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

Broda, C. and D. E. Weinstein (2010). Product creation and destruction: Evidence and price implications. *American Economic Review 100*, 691–723.

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.

Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with PyBLP. *RAND Journal of Economics 51*(4), 1108–1161.

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

Dubé, J.-P., , J. T. Fox, and C.-L. Su (2012). Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica 80*(5), 2231–2267.

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

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.

Frisch, R. (1933). *Pitfalls in the Statistical Construction of Demand and Supply Curves*. Leipzig: Hans Buske, Verlag.

Gandhi, A. and J.-F. Houde (2020). 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., W. K. Newey, and W. E. Taylor (1987). Efficient estimation and identification of simultaneous equation models with covariance restrictions. *Econometrica 55*, 1377–1398.

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.

Knittel, C. R. and K. Metaxoglou (2014). Estimation of random-coefficients demand models: Two empiricists' perspective. *The Review of Economics and Statistics 96*(1), 34–59.

Koopmans, T. C. (1949). Identification problems in economic model construction. *Econometrica 17*(2), 125–144.

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.

Leontief, W. (1929). Ein versuch zur statistichen analyse von angebot und nachfrage. *Weltvirtschaftliches Archiv 30*(1), 1–53.

Li, S., J. Mazur, Y. Park, J. Roberts, A. Sweeting, and J. Zhang (2021). Repositioning and market power after airline mergers.

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. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics & Management Strategy 9*(4), 513–548.

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.

Reynaert, M. and F. Verboven (2014). Improving the performance of random coefficients demand models: The role of optimal instruments. *Journal of Econometrics 179*(1), 83–98.

Romeo, C. (2010). Filling out the instrument set in mixed logit demand systems for aggregate data.

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.

Soderbery, A. (2015). Estimating import supply and demand elasticities: Analysis and implications. *Journal of International Economics 96*(1), 1–17.

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

Villas-Boas, S. B. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *Review of Economic Studies 74*(2), 625–652.

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. (2020). Dynamic airline pricing and seat availability.

Zoutman, F. T., E. Gavrilova, and A. O. Hopland (2018). Estimating both suppy and demand elasticities using variation in a single tax rate. *Econometrica 86*(2), 763–771.

A Demand System Applications and Extensions

The demand system of equation (1) 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 (1) takes the form

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

where $\overline{s}_{j|g} = \sum_{j \in \mathscr{J}_g} \frac{s_j}{\sum_{j \in \mathscr{J}_g} s_j}$ 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,

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

Thus, we verify that the derivatives can be expressed as a function of data and the non-linear parameters, allowing for three-stage estimation. In our third application, we use the nested logit model to estimate bounds on the structural parameters (Section 6.3).

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, 2020; Miller and Weinberg, 2017). We demonstrate how to estimate these parameters using additional covariance restrictions in our first application (Section 6.1).

The semi-linear demand assumption (Assumption 1) can be modified to allow for semi-linearity in a transformation of prices, $f(p_{jt})$:

$$h_{jt} \equiv h(q_{jt}, w_{jt}; \sigma) = \beta f(p_{jt}) + x'_{jt}\alpha + \xi_{jt} \tag{A.1}$$

Under this modification assumptions, it is possible to employ a method-of-moments approach to estimate the structural parameters. When $f(p_{jt}) = \ln p_{jt}$, the three-stage estimator and identification results are applicable, under the modified assumptions that ξ is orthogonal to $\ln X$ and that $\ln \eta$ and ξ are uncorrelated.

The optimal price for these demand systems takes the form $p_{jt} = \mu_{jt} c_{jt}$, where μ_{jt} is a markup that reflects demand parameters and (in general) demand shocks. It follows that the probability limit of an OLS regression of h on $\ln p$ is given by:

$$\beta^{OLS} = \beta - \frac{1}{\beta}\frac{Cov(\ln \mu, \xi)}{Var(\ln p^*)} + \frac{Cov(\ln \eta, \xi)}{Var(\ln p^*)}. \tag{A.2}$$

Therefore, the results developed in this paper are extend in a straightforward manner. We opt to focus on semi-linear demand throughout this paper for clarity.

A special case that is often estimated in empirical work is when h and $f(p)$ are logarithms:

Constant elasticity demand: With the modified demand assumption of equation (A.1), the constant elasticity of substitution (CES) demand model of Dixit and Stiglitz (1977) 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 and uncorrelatedness imply $Cov(p, \xi) = 0$, and OLS produces unbiased estimates of the demand parameters.[24] 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.

[24] The international trade literature following Feenstra (1994) consider non-constant marginal costs, which requires an additional restriction. See section 5.2 for an extension of our methodology to non-constant marginal costs.

The demand assumption in Equation (1) accommodates many rich demand systems. 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 $h(q_{jt}, w_{jt}; \sigma) \equiv q_{jt} - \sum_{k \neq j} \beta_{jk} p_k$. In addition to the own-product uncorrelatedness restrictions that could identify β_{jj}, one could impose cross-product covariance restrictions to identify β_{jk} ($j \neq k$). We discuss these cross-product covariance restrictions in Section 5.1. A similar approach could be used with the almost ideal demand system of Deaton and Muellbauer (1980).

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

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

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 (B.1). Thus, we eliminate the step of projecting p on x. This estimator is 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 holds in typical cases.

C 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

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

For each firm, stack the first-order conditions, writing the left-hand side as the product of a matrix A^m of loading components and a vector of markups, $(p_j - c_j)$, for products owned by the firm. The loading components are given by $A^m_{i(j),i(k)} = \frac{1}{dh_j/dq_k}$, 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:

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

Here, $(\Lambda^m \boldsymbol{q}^m)_{i(j)}$ 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 (11).[25] Letting $\tilde{h} \equiv (\Lambda^m \boldsymbol{q}^m)_{i(j)}$ 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 4. *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:*

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

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

[25] 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.

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:

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

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

Part 1: Equivalence

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

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

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:

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

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 1 (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

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

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 2 (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 (8) 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 (8), $a = 1$, 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* (8) *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 (8) 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

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

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

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

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

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

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

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

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

$$+ 4\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)} + 4\frac{1}{\beta}\frac{Cov\left(\xi, \frac{dh}{dq}q\right)}{Var(p^*)}\frac{Cov\left(p^*, \frac{dh}{dq}q\right)}{Var(p^*)}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 (8), 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.[26] 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

[26] 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).

because $\beta < 0$ so $-\beta > 0$. It follows that β is the lower root of (8) 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 Proposition 3 (Approximation)

The demand and supply equations are given by:

$$\begin{aligned} h =& \beta p + x'\alpha + \xi \\ p =& x'\gamma - \frac{1}{\beta}\frac{dh}{dq}q + \eta \end{aligned}$$

Using an first-order expansion of h about q, $h \approx \overline{h} + \overline{\frac{dh}{dq}}\left(q - \overline{q}\right)$, we can solve for a reduced-form for p and h. It follows that

$$\begin{aligned} \overline{h} + \overline{\frac{dh}{dq}}\left(q - \overline{q}\right) \approx& \beta p + x'\alpha + \xi \\ \overline{\frac{dh}{dq}} q \approx& \beta p + x'\alpha + \xi - \overline{h} + \overline{\frac{dh}{dq}}\overline{q} \end{aligned}$$

Letting $\frac{dh}{dq}q = \frac{\tilde{dh}}{dq}q + \overline{\frac{dh}{dq}}q$, we have

$$\begin{aligned} p &\approx x'\gamma - \frac{1}{\beta}\frac{\tilde{dh}}{dq}q - \frac{1}{\beta}\left(\beta p + x'\alpha + \xi - \overline{h} + \overline{\frac{dh}{dq}}\overline{q}\right) + \eta \\ 2p &\approx x'\gamma + \frac{1}{\beta}x'\alpha - \frac{1}{\beta}\overline{h} + \frac{1}{\beta}\overline{\frac{dh}{dq}}\overline{q} - \frac{1}{\beta}\frac{\tilde{dh}}{dq}q + \eta + \frac{1}{\beta}\xi \\ p &\approx \frac{1}{2}\left(x'\gamma + \frac{1}{\beta}x'\alpha - \frac{1}{\beta}\overline{h} + \frac{1}{\beta}\overline{\frac{dh}{dq}}\overline{q} - \frac{1}{\beta}\frac{\tilde{dh}}{dq}q + \eta + \frac{1}{\beta}\xi\right). \end{aligned}$$

Let H^* denote the residual from a regression of $\frac{\tilde{dh}}{dq}q$ on x. Then p^*, the residual from a regression of p on x, is

$$p^* \approx \frac{1}{2}\left(\eta + \frac{1}{\beta}\xi - \frac{1}{\beta}H^*\right). \tag{D.5}$$

Likewise, as $h - \overline{h} + \overline{\frac{dh}{dq}}\overline{q} \approx \overline{\frac{dh}{dq}}q$,

$$p \approx x'\gamma - \frac{1}{\beta}\frac{\tilde{dh}}{dq}q - \frac{1}{\beta}\overline{\frac{dh}{dq}}q + \eta$$

$$\begin{aligned}
h \approx& \beta \left(x'\gamma - \frac{1}{\beta}\frac{\tilde{dh}}{dq}q - \frac{1}{\beta}\frac{\overline{dh}}{dq}q + \eta \right) + x'\alpha + \xi \\
h \approx& \beta x'\gamma + x'\alpha - \frac{\tilde{dh}}{dq}q - \left(h - \overline{h} + \frac{\overline{dh}}{dq}\overline{q} \right) + \beta\eta + \xi \\
2h \approx& \beta x'\gamma + x'\alpha - \frac{\tilde{dh}}{dq}q + \overline{h} - \frac{\overline{dh}}{dq}\overline{q} + \beta\eta + \xi.
\end{aligned}$$

Similarly, the residual from a regression of h on x is:

$$h^* \approx \frac{1}{2}\left(\beta\eta + \xi - H^*\right). \tag{D.6}$$

Equations (D.5) and (D.6) provide an approximation for β.

$$\begin{aligned}
-\sqrt{\frac{Var(h^*)}{Var(p^*)}} &\approx -\sqrt{\frac{\frac{1}{4}Var\left(\beta\eta + \xi - H^*\right)}{\frac{1}{4}Var\left(\eta + \frac{1}{\beta}\xi - \frac{1}{\beta}H^*\right)}} \\
&\approx -\sqrt{\frac{\beta^2 Var\left(\eta + \frac{1}{\beta}\xi - \frac{1}{\beta}H^*\right)}{Var\left(\eta + \frac{1}{\beta}\xi - \frac{1}{\beta}H^*\right)}} \\
&\approx \beta
\end{aligned}$$

QED.

Proof of Lemma 2 (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, $a = 1$.

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

$$1 + \frac{b}{(b^2 - 4c)^{\frac{1}{2}}} > 0. \tag{D.7}$$

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.7). Second, when $c > 0$, it must be the case that $b > 0$ also. Otherwise, both roots are positive, invalidating the model. When $b > 0$, it is evident that the left-hand side of (D.7) 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, $U \equiv \frac{1}{2}\left(-b + \sqrt{b^2 - 4ac}\right)$ is 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 4 (Covariance Bound)

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 $b^2 \geq 4ac$. Given the formulation of (8), real solutions satisfy the condition:

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

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

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

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

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

This bound exists if the expression inside the radical is positive, which is the case if and only if the sufficient condition for point identification from Proposition 2 fails. QED.

Alternative Proof

Alternatively, one can calculate the minimum of equation (23), which is repeated here for convenience.

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

The minimum is obtained at

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

which has the solution

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

The sign of the solution is determined because β must be negative. Verifying the second-order condition for a minimum is straightforward. Therefore, the covariance bound may be expressed as:

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

Proof of Proposition 5 (Non-Constant Marginal Costs)

Under the semi-linear marginal cost schedule of equation (16), 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

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

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 Proposition B.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

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

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.