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

Alexander MacKay Harvard University†

Nathan H. Miller Georgetown University‡

August 29, 2023

Abstract

We consider the identification of empirical models of supply and demand with imperfect competition. We show that a covariance restriction on unobserved demand and cost shocks resolves endogeneity and identifies the price parameter. We demonstrate how to employ this approach in estimation, and we provide a comparison to instrumental variables approaches. Our formal results also indicate how weaker assumptions about the covariance term can be used to construct bounds on the price parameter. We illustrate the covariance restriction approach 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 Chris Adams, Steven Berry, Renata Gaineddenova, Charlie Murry, Matt Osborne, Chuck Romeo, Gloria Sheu, Karl Schurter, Jesse Shapiro, Michael Sullivan, 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 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 using the supply model to recover marginal costs and simulate counterfactuals (e.g., Berry et al., 1995; Nevo, 2001).

In this paper, we develop an alternative identification strategy that exploits covariance restrictions between demand-side and supply-side structural error terms. We first establish a function linking the price parameter and the covariance of unobservable cost and demand shocks for a broad class of oligopoly models. We then show how this relationship can identify the price parameter and pin down the slope of demand in estimation. A key distinction between our approach and the use of instrumental variables is that we interpret endogenous variation in prices and quantities through the lens of the model, rather than relying on an additional (observed) variable to isolate exogenous variation in price. The core intuition is that the supply side of the model dictates how prices respond to demand shocks, shaping the relative variation of quantities and prices in the data. We explore the promise and limitations of the covariance restrictions approach to estimation, both theoretically and in the context of three empirical applications that we draw from the literature.

In Section 2, we outline the data generating process for our baseline model and provide formal identification results. The model can accomodate standard empirical demand systems such as logit and random coefficients logit, among others. The supply-side assumptions nest different models of conduct for oligopolists with constant marginal costs, including differentiated-products Bertrand competition and Cournot competition. In this setting, prices are endogenous because they respond to a demand shock (the demand-side "structural error term") that is unobserved to the econometrician.

We prove that the price parameter solves a quadratic equation in which the coefficients are functions of observables and the covariance of demand and cost shocks. With a restriction on the covariance term, the price parameter is identified up to (at most) two points. Under reasonable conditions, the price parameter is the more negative solution, and point identification is obtained. The price parameter can be computed directly from an analytical solution, or the covariance restriction can be recast as an orthogonality condition and estimation can proceed with the method of moments. We show how the empirical variation in (transformed) quantities and prices is informative about the price parameter. Intuitively, the greater the relative variation in quantities, the more elastic demand must be, all else equal.[1]

[1]This is true even if the empirical relationship between prices and quantities is upward sloping. Our method will still recover the correct downward-sloping demand curve.

Point identification can thus be achieved with a fully-specified model and a covariance restriction on unobserved shocks. Whether the covariance restriction is credible depends on details of the economic environment. For example, in the presence of capacity constraints, a positive demand shock can increase marginal costs, yielding an overall positive correlation. Using weaker assumptions, we show how our results can be used achieve partial identification and bound the price parameter. If the econometrician can sign the correlation between unobserved demand and cost shocks, then one-sided bounds can be placed on the price parameter. Furthermore, our results show when it is possible to rule out some values of the price parameter without any assumption on this correlation.

Like instrumental variables, covariance restrictions rely on an assessment about the distribution of unobservables. In Section 3, we make this comparison explicit by comparing the assumptions needed to obtain point identification with either approach. As typically defined, valid instruments satisfy an exclusion condition and a relevance condition. The exclusion restriction for instrumental variables is similar to the covariance restriction when the econometrician uses cost variation as an instrument. However, the instrumental variables approach requires the econometrician to observe exogenous variation in order to identify the model (e.g., Wooldridge, 2010; Berry and Haile, 2014). The covariance restriction approach avoids this requirement by imposing the stronger assumption that the residual unobserved variation is uncorrelated. This assumption also precludes the need for a relevance condition; all of the observed variation in prices and quantities is used in estimation. By contrast, the instrumental variables approach can suffer from bias if the instrument and price are weakly correlated.

After formalizing these distinctions, we compare our identification strategy to an approach that uses "residual instruments" recovered under a covariance restriction (Hausman and Taylor, 1983) and to the oft-used instruments of Hausman (1996). Finally, we use Monte Carlo simulations to show that, when the needed assumptions are met, a covariance restriction outperforms instruments in finite samples, especially when instruments become weak.

In Section 4, we consider extensions to the baseline model. First, we consider a more general class of covariance restrictions that might be employed in practice. For example, for the Hausman instruments—prices in related markets—to be valid, three economic assumptions about the correlation structure between unobserved shocks must be satisfied. We discuss how one could employ these assumptions directly in a method-of-moments estimator, rather than relying on an instrumental variables implementation using observed prices.

We also discuss models in which costs are not constant in quantity. In such cases, the response of prices to demand shocks is mediated by the slope of the cost curve. One approach to resolve this issue is to explicitly model the (non-constant) marginal cost function. A covariance restriction may then be credible, though identification requires additional moments for any parameters that enter the non-constant portion of marginal costs. Alternatively, one could forgo the estimation of the marginal cost function and instead invoke the bounds approach.

We apply these methods in a series of empirical applications (Section 5). Each of the three settings that we have selected—ready-to-eat (RTE) cereals, cement, and airlines—differ in a variety of ways that influence our implementation. With RTE cereals, marginal costs can plausibly be modeled as constant, so we proceed with estimation under a covariance restriction, using fixed effects to absorb potentially confounding variation. With cement, capacity constraints imply that marginal costs may increase with quantities. We follow an approach developed in the literature and model this effect explicitly, after which we view a covariance restriction as credible. Finally, with airlines, we apply a bounds approach. In each case, we show how covariance restrictions support inferences about the price parameter.

Together, these results indicate that covariance restrictions can help overcome a significant obstacle for empirical research—finding valid instruments for price—in some environments. In our three applications, we obtain parameter values that are consistent with instrumental variable estimates. Outside of this paper, Döpper et al. (2023) show that covariance restrictions deliver estimates comparable to those in the industrial organization literature (for cereal, yogurt, and beer) as well the international trade literature (coffee). De los Santos et al. (2021) use covariance restrictions in the context of e-books. Like other electronic goods, e-books can have substantial fluctuations in demand over time that are unrelated to changes in marginal costs. These examples indicate that covariance restrictions can deliver reasonable estimates in different settings.

Empirical models of imperfect competition typically have other key parameters that characterize heterogeneity of consumer preferences. To identify these parameters, researchers have used micro-moments constructed from the observed behavior of individual consumers (e.g., Backus et al., 2021; Döpper et al., 2023) or "second-choice" data on what consumers view as their next-best option (e.g., Grieco et al., 2023). These strategies identify the consumer heterogeneity parameters but do not resolve price endogeneity (Berry and Haile, 2020). Thus, the covariance restriction we examine is a useful complement to the use of detailed consumer data.[2]

To put our results in context, covariance restrictions were analyzed in early research at the Cowles Foundation on the identification of linear systems of equations, including supply and demand models of perfect competition (e.g., Koopmans, 1949; Koopmans et al., 1950).[3] With perfect competition, the supply curve is upward-sloping due to increasing costs of production. With upward-sloping supply and downward-sloping demand, two separate restrictions are required for identification (Hausman and Taylor, 1983). If instead price endogeneity arises due to the markup adjustments that occur in models of imperfect competition, then (as we show)

[2]Alternatively, if instruments constructed from the characteristics of competing products (e.g., Berry et al., 1995; Gandhi and Houde, 2020), then the covariance restriction could be incorporated using the generalized method of moments (GMM) as an additional identifying restriction.

[3]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.

only a single restriction is sufficient for identification.

The strategy of using supply-side restrictions to reduce identification requirements has parallels in a handful of other articles. 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 constant-markup approach (e.g., Broda and Weinstein, 2006, 2010; Soderbery, 2015).[4] 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. At a high level, our approach to estimation with covariance restrictions relates to Petterson et al. (2022), which shows how to bound structural parameters based on beliefs about the magnitudes of unobserved shocks. Our research builds on these articles by developing results for imperfect competition with adjustable markups.[5]

2 Model and Identification

2.1 Data Generating Process

The model examines supply and demand across markets. Markets can be conceptualized as (for example) separate geographies, time periods, or both. In each market t, there is a set $\mathcal{J}_t = \{0, 1, \ldots, J_t\}$ of products available for purchase. The market $t = 1, \ldots, T$ is defined by $(\mathcal{J}_t, \chi_t)$, where

$$\chi_t = \{\boldsymbol{x}_t, \boldsymbol{w}_t, \boldsymbol{\xi}_t, \boldsymbol{\eta}_t\}$$

contains product and market characteristics. Among these, $\boldsymbol{x}_t = [x_{1t}; \ldots; x_{J_t t}]$ is a $J_t \times K$ matrix of (non-price, exogeneous) product-market characteristics that are observable to the econometrician, $\boldsymbol{w}_t$ contains observable variables that arise in some specifications of the model, and $\boldsymbol{\xi}_t = (\xi_{1t}, \ldots, \xi_{J_t t})$ and $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{J_t t})$ are mean-zero $J_t \times 1$ vectors of unobservable product-level or market-level characteristics. We sometimes refer to the unobservable characteristics as "structural error terms." Let each $\xi_{jt}, \eta_{jt} \in \mathbb{R}$ and each x_{jt} be a $K \times 1$ (row) vector $\in \mathbb{R}^K$. We assume that the first element in each x_{jt} equals one, i.e., that the characteristics in $\boldsymbol{x}_t$ include a constant. The dimension of $\boldsymbol{w}_t$ depends on the modeling specification. Without loss of generality, let $\mathcal{J}_t = \mathcal{J} = \{0, 1, \ldots, J\}$ going forward.

Prices and quantities are determined endogenously by market participants. Let $\boldsymbol{p}_t = (p_{1t}, \ldots, p_{Jt})$

[4] 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.

[5] Some applications in industrial organization identify demand-side parameters with the assistance of supply-side assumptions (e.g., Thomadsen, 2005; Cho et al., 2018; Li et al., 2022). Among these, Thomadsen assumes there are not unobserved demand shocks and Cho et al. assume that there are no unobserved cost shocks; both are special cases of the covariance restriction approach.

be a vector of prices and $\boldsymbol{q}_t = (q_{1t}, \ldots, q_{Jt})$ be a vector of quantities, with $p_{jt}, q_{jt} \in \mathbb{R}$. Both prices and quantities are observable to the econometrician. The parameters of the model are in the set $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$. Following Nevo (2001), we let θ_1 include parameters that affect demand and supply linearly in a manner that we specify below, whereas we use θ_2 for additional parameters that enter with some specifications of the model. An example of the latter is the nesting parameter that enters if demand is nested logit (e.g., Berry, 1994). Our main identification results are for θ_1. Covariance restrictions also can help pin down θ_2 in some settings.

On the demand side, we assume that the quantity of each product is determined by $q_{jt} = \sigma^{(j,t)}(\boldsymbol{p}_t, \chi_t; \boldsymbol{\theta})$, where each $\sigma^{(j,t)}$ is a demand function. We also assume that, for every (j,t), there exists a known function $h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2)$ that is increasing in own quantity (q_{jt}). The function can be interpreted as providing transformed quantities as a function of $\boldsymbol{w}_t$ and θ_2. We provide examples of $h^{(j,t)}$ later as its form depends on the demand system. The substantive restriction we place on demand is that $h^{(j,t)}$ is constructed such that the following equality is satisfied everywhere:

$$h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) = \alpha p_{jt} + x_{jt}\beta + \xi_{jt} \tag{1}$$

where α and β are parameters contained in θ_1 and $\alpha < 0$ (i.e., demand slopes down). Thus, we assume that a known function can map quantities into an index that is linear in prices. Models that satisfy these conditions are used regularly in the empirical industrial organization literature.[6]

This restriction embeds two important assumptions. The first is that the unknowns θ_1, $\boldsymbol{\xi}_t$, and $\boldsymbol{\eta}_t$ do not enter $h^{(j,t)}$ directly (but can indirectly through $\boldsymbol{q}_t$). Thus, $h^{(j,t)}$ can be constructed given observables and knowledge of θ_2. The second is that the right-hand side of equation (1) is linear. This restriction allows us to use linear regression results to construct an analytic expression for α. Though separability of ξ_{jt} and p_{jt} is important, it is not critical that the expression be linear in prices, as we show in Appendix A.

On the supply side, we decompose prices into markups and marginal costs:

$$p_{jt} = \mu_{jt} + mc_{jt} \tag{2}$$

Consistent with equilibrium behavior in a broad class of oligopoly models, we assume that, for each (j,t), there exists a known function $\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2)$ such that the following equation holds:

$$\mu_{jt} = -\frac{1}{\alpha}\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) \tag{3}$$

The substantive restrictions imposed are the multiplicative separability of $\frac{1}{\alpha}$ and that the un-

[6]See the discussion of logit demand below. Nested logit, random coefficients logit, linear demand, and constant elasticity demand are nested within the general model or accommodated with straightforward generalizations (Appendix A).

knowns θ_1, $\boldsymbol{\xi}_t$, and $\boldsymbol{\eta}_t$ do not enter $\lambda^{(j,t)}$ directly (but, as with demand, they can enter indirectly through $\boldsymbol{q}_t$). Therefore, $\lambda^{(j,t)}$ can be constructed given observables and knowledge of θ_2. We show how to construct $\lambda^{(j,t)}(\cdot)$ in a variety of specific contexts in our applications and in Appendix A.[7]

To fix ideas, consider the canonical model of logit demand with oligopoly price competition. Quantities are given by $q_{jt} = s_{jt}M_t$ where s_{jt} is the market share of the product and M_t (contained in $\boldsymbol{w}_t$) is the size of market t. The left-hand side of equation (1) is constructed as $h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) \equiv \ln(s_{jt}) - \ln(s_{0t})$, where $s_{0t} = 1 - \sum_{k\in\mathcal{J}} s_{kt}$ is the market share of the "outside good." With logit demand, there are no parameters in θ_2, and market size is the only variable in $\boldsymbol{w}_t$. The $h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2)$ function provides the utility that the average consumer would obtain from the product. Likewise, on the supply side, $\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2)$ provides the markups *measured in utils*, while dividing by $-\alpha$ obtains markups measured in units of currency. With single product firms, the logit markup is given by $\mu_{jt} = -\frac{1}{\alpha}\frac{1}{1-s_{jt}}$ and thus $\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) = \frac{1}{1-s_{jt}}$.

We initially maintain that marginal costs are constant and linear in the product characteristics:

$$mc_{jt} = x_{jt}\gamma + \eta_{jt} \tag{4}$$

where γ is contained in θ_1. We later allow marginal costs to depend on quantities (Section 4.2).

Combining equations (2)-(4), the supply side of the model implies that the following equation is satisfied for each product j and market t:

$$\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) = -\alpha p_{jt} + \alpha x_{jt}\gamma + \alpha\eta_{jt} \tag{5}$$

This *supply relationship* characterizes how prices and quantities respond to shifts in demand (holding fixed α and marginal costs) given the behavior of firms. It can, for example, capture optimal price-setting behavior when individual firms have market power, and, unlike the supply curve for perfectly competitive firms, the supply relationship can lie above the marginal cost curve when plotted in price-quantity space.[8] Because the supply relationship expresses (transformed) quantities as a linear function of prices and characteristics, it is the analog to the demand relationship of equation (1).

Together, equations (1) and (5) provide the conditions that jointly determine prices and quantities. The supply-side behavior captured by equation (5) does not necessarily have to correspond to equilibrium behavior, but, when it does, these equations yield equilibrium outcomes.

[7]For example, with single-product Bertrand pricing and differentiable demand, we have the general expression $\mu_{jt} = -\frac{1}{dq_{jt}/dp_{jt}}q_{jt} = -\frac{1}{\alpha}\frac{dh^{(j,t)}}{dq_{jt}}q_{jt}$, yielding $\lambda^{(j,t)} = \frac{dh^{(j,t)}}{dq_{jt}}q_{jt}$. Appendix A provides the relevant forms for specific demand systems, addresses the construction of $\lambda^{(j,t)}$ with multi-product firms, and covers a generalized model of oligopoly that nests both Bertrand competition in prices and Cournot competition in quantities.

[8]The difference between marginal costs and the supply relationship is the (perceived) inframarginal loss in revenue for selling an additional unit of quantity. Bresnahan (1982) refers to the inverse of this equation, with price on the left-hand side, as the "supply relation" and notes that it generalizes to different models of firm conduct. See Appendix A.6 for a figure and additional discussion.

The framework covers many of the empirical models of industrial organization (Appendix A). Nonetheless, some models are excluded. For example, in models with constant elasticity demand, one cannot construct $h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2)$ that satisfies equation (1). It is possible, however, to find an function that satisfies a related restriction: that the right-hand side is linear in log prices and log prices are additively separable from the demand-side structural error term. We discuss how to extend our results to this and other cases in Appendix A.

Equations (1) and (5) also illuminate potential strategies to identify θ_1. If a characteristic x^k shifts marginal costs ($\gamma^k \neq 0$) but is excluded from demand ($\beta^k = 0$) then it is a valid instrument and can be used to estimate equation (1). A variable not in x that is correlated with η but not ξ could also be a valid instrument. Conversely, a variable that shifts demand but is excluded from marginal costs can be a valid instrument in the estimation of equation (5). Thus, both cost shifters and demand shifters can provide exogenous variation that identifies θ_1. Similarly, "markup shifters" that create variation in $\lambda^{(j,t)}$ or $h^{(j,t)}$ can be valid instruments for either the supply-side or the demand-side.[9] Another possibility is to place a covariance restriction between the structural error terms from both equations, which we develop below.

2.2 Identification with Covariance Restrictions

We now consider the identification of the model using a covariance restriction, focusing on the linear parameters $(\alpha, \beta, \gamma) = \theta_1$. Therefore, we assume that the econometrician knows θ_2. The identification of θ_2 has been considered in other research (e.g., Berry et al., 1995; Berry and Haile, 2020; Gandhi and Houde, 2020), and we return to the prospect that covariance restricts may identify θ_2 in the context of the empirical applications.

Stacking objects across markets, the econometrician observes vectors of prices and quantities ($\boldsymbol{P}$ and $\boldsymbol{Q}$), a matrix of non-price characteristics ($\boldsymbol{X}$), and (possibly) other observables ($\boldsymbol{W}$). Using observables, the model, and θ_2, the econometrician can evaluate the demand and supply transformations $h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2)$ and $\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2)$ for every (j, t) pair. We denote these values h_{jt} and λ_{jt} and treat them as observed. The stacked JT vector for h_{jt} we denote $\boldsymbol{H}$.

The structural error terms can be decomposed as follows:

$$\xi_{jt} = \xi_j + \xi_t + \Delta\xi_{jt} \tag{6}$$

$$\eta_{jt} = \eta_j + \eta_t + \Delta\eta_{jt} \tag{7}$$

which incorporates product-specific persistent components (e.g., higher quality or higher cost), market-specific components (greater demand in a year and/or region), and an orthogonal mean-zero residual term. Define an augmented characteristics matrix, $\tilde{\boldsymbol{x}}_t$, as including the

[9] These might include functions of other products' characteristics (Berry et al., 1995; Gandhi and Nevo, 2021) or competitive events such as mergers, entry, or exit (e.g., Miller and Weinberg, 2017).

K observed covariates and a full set of dummy variables for products $(J-1)$ and markets $(T-1)$. Stacking across markets, we obtain $\tilde{\boldsymbol{X}}$ with dimension $JT \times (K+J+T-2)$.

We assume the augmented characteristics are exogenous in the sense that $\mathbb{E}[\Delta\xi_{jt}|\tilde{\boldsymbol{x}}_t] = \mathbb{E}[\Delta\eta_{jt}|\tilde{\boldsymbol{x}}_t] = 0$ for all $j = 1, \ldots, J$, as is commonly maintained in the literature. We also assume that $\mathbb{E}[\boldsymbol{v}_t'\boldsymbol{v}_t]$ has full rank, where $\boldsymbol{v}_t = [\boldsymbol{p}_t \ \tilde{\boldsymbol{x}}_t]$ is a $J \times (K+J+T-1)$ matrix that combines prices with the augmented characteristics. Two immediate implications of these assumptions are that $\mathbb{E}[\tilde{\boldsymbol{x}}_t'\tilde{\boldsymbol{x}}_t]$ has full rank and that β and γ are trivially identified given knowledge of α, following the standard arguments for linear regression.

We focus on the identification of the price parameter, α, in the remainder of this section. We leverage a single moment (a covariance restriction), which allows us to express our results using variance and covariance terms among random variables. This notation is familiar from univariate regressions. We obtain univariate regression analogs by taking a single variable (p_{jt}), projecting it on the other characteristics $(\tilde{\boldsymbol{x}}_{jt})$, and then considering a regression with the residualized values as the single regressor. By the Frisch-Waugh-Lovell theorem (and exogeneity of $\tilde{\boldsymbol{x}}_t$), the resulting coefficient estimate is identical to one obtained in the full multivariate regression. The residuals from a regression of p_{jt} on $\tilde{x}_{jt}$ are given by

$$\boldsymbol{P}^* = \boldsymbol{P} - \tilde{\boldsymbol{X}}[\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}}]^{-1}\tilde{\boldsymbol{X}}'\boldsymbol{P}, \tag{8}$$

and they provide the component of price that is orthogonal to characteristics and fixed effects. Later in this paper we residualize other variables in the same fashion. Throughout, we will use the superscript * to denote the residuals obtained from a regression of a variable on $\tilde{x}_{jt}$.

An implication of the rank condition is that the augmented characteristics do not fully explain prices. Therefore, the unconditional variance of p^*_{jt} is positive, i.e., $Var(p^*) > 0$.[10] Here and throughout the remainder of the paper, we omit jt subscripts when the variable occurs in a covariance or variance expression or in the body of the text, e.g., p refers to p_{jt} and x refers to x_{jt}. Note that p^* is distinguished from the length JT vector $\boldsymbol{P}^*$ by the use of lower case and a lack of bold font.

We now formalize our first identification result, which links the OLS estimate to the price parameter. Following the discussion above, the probability limit ($T \to \infty$) of the OLS estimate of α obtained from a regression of h on p and $\tilde{x}$ is

$$\alpha^{OLS} \equiv \frac{Cov(p^*, h)}{Var(p^*)} = \alpha + \frac{Cov(p^*, \Delta\xi)}{Var(p^*)}. \tag{9}$$

The corresponding OLS residuals are given by $\boldsymbol{\Delta\xi}^{OLS} = \boldsymbol{H} - \boldsymbol{V}[\boldsymbol{V}'\boldsymbol{V}]^{-1}\boldsymbol{V}'\boldsymbol{H}$.

We now construct a function that maps the price coefficient to a specific value for the co-

[10] This expression refers to the unconditional variance of p^* (over products and markets). The unconditional variance is defined as $Var(p^*) \equiv \mathbb{E}[(p^*)^2] - \mathbb{E}[p^*]^2$ where the expectations are taken over markets and products. The second component is zero because we assume that x includes a constant. The empirical analog is $\frac{1}{JT}\sum_{t=1}^{T}\sum_{j=1}^{J}(p^*_{jt})^2$.

variance of the residual structural error terms:

Proposition 1. *The probability limit of the OLS estimate can be written as a function of α, the residuals from an OLS regression, prices and quantities, and a covariance term:*

$$\alpha^{OLS} = \alpha - \frac{1}{\alpha + \frac{Cov(p^*,\lambda)}{Var(p^*)}} \frac{Cov\left(\Delta\xi^{OLS}, \lambda\right)}{Var(p^*)} + \alpha \frac{1}{\alpha + \frac{Cov(p^*,\lambda)}{Var(p^*)}} \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} \tag{10}$$

Therefore, α solves the following quadratic equation:

$$\begin{aligned} 0 = \; & \alpha^2 \\ & + \left(\frac{Cov\left(p^*, \lambda\right)}{Var(p^*)} - \alpha^{OLS} + \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} \right) \alpha \\ & + \left(-\alpha^{OLS} \frac{Cov\left(p^*, \lambda\right)}{Var(p^*)} - \frac{Cov\left(\Delta\xi^{OLS}, \lambda\right)}{Var(p^*)} \right) \end{aligned} \tag{11}$$

All proofs are in Appendix D. The terms in equation (11) are well defined under our rank condition and, aside from α and $Cov(\Delta\xi, \Delta\eta)$, they have straightforward empirical analogs.

There are at most two solutions for α for any given value of $Cov(\Delta\xi, \Delta\eta)$. Further, in most empirical models, α is likely to be the lower root. The following result provides formal conditions under which this is guaranteed:

Proposition 2. *The parameter α is the lower root of equation* (11) *if and only if*

$$-\frac{1}{\alpha} Cov(\Delta\xi, \Delta\eta) \leq Cov\left(p^*, \Delta\eta - \frac{1}{\alpha}\Delta\xi\right) \tag{12}$$

and, furthermore, α is the lower root of equation (11) *if*

$$0 \leq \alpha^{OLS} Cov\left(p^*, \lambda\right) + Cov\left(\Delta\xi^{OLS}, \lambda\right) \tag{13}$$

In the first condition, it is helpful to think of $-\frac{1}{\alpha}\Delta\xi$ as the residual demand-side structural error term, scaled so that units are equivalent to those of marginal costs (and price). If $Cov(\Delta\xi, \Delta\eta) = 0$, the condition holds as long as prices tend to increase with demand and marginal costs, as occurs in most empirical models. For example, the condition holds when demand is linear. Thus, α is likely the lower root of equation (11) in most applications.

The second condition is derived using properties of the quadratic formula. Because the terms in equation (13) are constructed from data, the sufficient condition can be estimated and used to test (and possibly reject) the null hypothesis that multiple negative roots exist. Henceforth, we assume that α is the lower root of equation (11).

The implication of this result—a one-to-one function mapping α to $Cov(\Delta\xi, \Delta\eta)$—is that the price coefficient can be recovered with information about the correlation between residual demand and cost shocks in models with imperfect competition. Conversely, moments that pin down the price parameter also pin down the value of $Cov(\Delta\xi, \Delta\eta)$.

2.3 Estimation with Covariance Restrictions

Estimation can proceed with the method of moments (a general approach) by recasting the information about the covariance term as an orthogonality condition. One possibility is that that demand-side and supply-side structural error terms are uncorrelated: $Cov(\Delta\xi, \Delta\eta) = 0$. Equivalently, this can be expressed as $\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{jt}] = 0$.

Under this assumption, the method-of-moments estimator uses the empirical analog of this condition and attempts to minimize its contribution to the objective function. For a case with one moment and one parameter, the method-of-moments estimate of α is given by

$$\hat{\alpha}^{CR} = \arg\min_{\tilde{\alpha}<0} \left[\frac{1}{T}\frac{1}{J} \sum_{t} \sum_{j\in\mathcal{J}} \Delta\xi_{jt}(\tilde{\alpha})\Delta\eta_{jt}(\tilde{\alpha}) \right]^2, \tag{14}$$

where $\Delta\xi_{jt}(\tilde{\alpha})$ and $\Delta\eta_{jt}(\tilde{\alpha})$ can be recovered from residualized (transformed) quantities and prices given the candidate parameter under consideration. Some care must be taken to ensure convergence to the lower root. The generalized method of moments (GMM) may also be used with additional moments or when estimating multiple parameters jointly, in which case the sample moment may be weighted against other components of the GMM objective function using the standard approach.

An implication of Proposition 1 is that this estimate is consistent for the price parameter, i.e., $\hat{\alpha}^{CR} \rightarrow \alpha$. This is notable because, in general, the inclusion of a moment in a method-of-moments approach does not imply the consistent identification of an additional parameter.[11] By contrast, a covariance restriction between the structural error terms provides identification of the price parameter *a priori*. Pushing the result further, one could use the analytical expression in equation (11) to directly compute the coefficient estimate.[12]

This method-of-moments approach is employed in our first application and also by Döpper et al. (2023) to estimate models with random-coefficients logit demand and Bertrand pricing. In these models, β and γ can be estimated with OLS regression once $\hat{\alpha}^{CR}$ is obtained. Döpper et al.

[11] To highlight this, consider that Berry and Haile (2020) identify a class of moment conditions (micro-moments) that can pin down consumer heterogeneity but provide *no identifying information* about the price parameter.

[12] The probability limit of the coefficient estimate is given by:

$$\alpha^{CR} = \frac{1}{2}\left(\alpha^{OLS} - \frac{Cov\,(p^*, \lambda)}{Var(p^*)} - \sqrt{\left(\alpha^{OLS} + \frac{Cov\,(p^*, \lambda)}{Var(p^*)}\right)^2 + 4\frac{Cov\,(\Delta\xi^{OLS}, \lambda)}{Var(p^*)}}\right), \tag{15}$$

which obtains from an application of the quadratic formula.

(2023) also illustrate how the additional parameters in θ_2 can be estimated with GMM using the nested fixed point approach of Berry et al. (1995). For each candidate $\tilde{\theta}_2$, the covariance restrictions estimator is applied to obtain $\hat{\alpha}^{CR}(\tilde{\theta}_2)$. In the outer loop, an estimate of θ_2 is pinned down by micro-moments in the GMM objective function. More generally, covariance restriction can be applied in conjunction with instruments, and additional moments allow for efficiency improvements and specification tests (e.g., Hausman, 1978; Hansen, 1982).

The empirical variation that identifies α is the relative variation of (transformed) quantities and prices. When $Cov(\Delta\xi, \Delta\eta) = 0$, we obtain the following formal result:

Proposition 3. *If $Cov(\Delta\xi, \Delta\eta) = 0$, then a first-order approximation to probability limit of the method-of-moments estimator is*

$$\alpha^{CR} \approx -\sqrt{\frac{Var(h^*)}{Var(p^*)}}. \tag{16}$$

Intuition can be gleaned from the simultaneous equations representation of the model using equations (1) and (5). Rearranging these to obtain inverse demand and inverse supply relationships, we have:

$$p_{jt} = \frac{1}{\alpha} h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) - \frac{1}{\alpha} x_{jt}\beta - \frac{1}{\alpha}\xi_{jt} \qquad \text{(Demand)} \tag{17a}$$

$$p_{jt} = -\frac{1}{\alpha}\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) + x_{jt}\gamma + \eta_{jt} \qquad \text{(Supply)} \tag{17b}$$

By inspection, α determines the slope of both equations. A large α corresponds to a flatter inverse demand schedule (i.e., price sensitive consumers) and a flatter inverse supply relationship (i.e., less market power). Uncorrelated shifts in such schedules tend to generate more variation in quantity than price. By contrast, a small α corresponds to steeper inverse demand and inverse supply, such that uncorrelated shifts generate more variation in price than quantity. Connecting these observations formally generates an approximation of the lower root based on the ratio of variances. We illustrate this argument using a numerical example in Appendix B.

2.4 Partial Identification: Bounds

We now show how our formal identification results can be used to construct bounds on the price parameter, which may be useful for inference when a covariance restriction along the lines of $Cov(\Delta\xi, \Delta\eta) = 0$ is not plausible. We first consider bounds that utilize prior knowledge of the sign of the correlation between $\Delta\xi$ and $\Delta\eta$. Next, we show how the model and the data together may bound the price coefficient without any additional information.

For the first case, we assume that the econometrician can sign the correlation between $\Delta\xi$ and $\Delta\eta$. This situation might arise, for example, if factor prices are influenced by macroeconomic conditions, such that there is a link between the unobserved demand-side and supply-side error terms that is difficult to model explicitly. With a prior of the sign of $Cov(\Delta\xi, \Delta\eta)$, bounds can be placed on α. The reason is that there is a one-to-one mapping between the value of $Cov(\Delta\xi, \Delta\eta)$ and the lower root of equation (11):

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

Thus, if idiosyncratic demand and costs are correlated, such as through capacity constraints ($Cov(\Delta\xi, \Delta\eta) \geq 0$), then one-sided bounds can be placed on α. More generally, let $r(m)$ be the lower root of the quadratic in equation (11), evaluated at $Cov(\Delta\xi, \Delta\eta) = m$. Then $Cov(\Delta\xi, \Delta\eta) \geq m$ produces $\alpha \in (-\infty, r(m)]$, and $Cov(\Delta\xi, \Delta\eta) \leq m$ produces $\alpha \in [r(m), 0)$. The lower root, $r(m)$, can be estimated with the method-of-moments.[13]

For the second case, it can be that some values of the price parameter are unable to rationalize the data for *any* amount of correlation between $\Delta\xi$ and $\Delta\eta$. These values can be ruled out. Thus, the demand and supply assumptions alone may be informative about the plausible range of α. Formally, this occurs when the quadratic from equation (11) does not have a lower root, and thus no valid solution for α. To see how this can occur, represent the quadratic from equation (11) as $az^2 + bz + c$. By assumption, one root is $\alpha < 0$. As $a = 1$, 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. This places restrictions on $Cov(\Delta\xi, \Delta\eta)$, which is a component of b. From the monotonicity result (Lemma 1), we can use the excluded values of $Cov(\Delta\xi, \Delta\eta)$ from this result to rule out values of α.

We now state the result formally:

Proposition 4. ***(Model-Based Bound)*** *The model and data alone may bound* $Cov(\Delta\xi, \Delta\eta)$ *from below. The bound is given by:*

$$Cov(\Delta\xi, \Delta\eta) > Var(p^*)\alpha^{OLS} - Cov(p^*, \lambda) + 2Var(p^*)\sqrt{\left(-\alpha^{OLS}\frac{Cov(p^*, \lambda)}{Var(p^*)} - \frac{Cov\left(\Delta\xi^{OLS}, \lambda\right)}{Var(p^*)}\right)}$$

The bound exists if and only if the term inside the radical is non-negative. Further, through equation (11)*, this lower bound on* $Cov(\Delta\xi, \Delta\eta)$ *provides an upper bound on* α.

From the monotonicity result (Lemma 1), the excluded values of $Cov(\Delta\xi, \Delta\eta)$ rule out values of α. A model-based upper bound for α is obtained by evaluating the lower root of equation (11) at the model-based bound of $Cov(\Delta\xi, \Delta\eta)$. In practice, priors over the the covariance of unobserved shocks may be combined with model-based bounds to further restrict the identified set.

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

2.5 Discussion

The covariance restriction approach to estimation requires the econometrician to assess whether a covariance restriction between the structural error terms of the model is reasonable given the institutional details of the setting. It also requires the econometrician to specify the supply-side in order to estimate demand parameters. We discuss these two requirements in this section.

There are settings in which an assumption that demand and cost shocks are uncorrelated would be problematic. First, when products vary in quality, we would typically expect that products with higher (unobserved) quality are produced at higher costs. If the econometrician only has the use of cross-sectional data, then fixed effects (ξ_j, η_j) cannot be used to absorb the confounding variation, and the estimates would be biased. A second setting that would generate correlation in demand and supply shocks is one in which marginal costs vary with quantity produced. A third situation is simply when, even after addressing the above issues, the econometrician anticipates that the residual shocks will be correlated.

In the above circumstances, an understanding that the structural error terms are likely correlated may also correspond to a belief that the correlation has a particular sign, or, potentially, falls in a range of values. In such, cases, the econometrician can pair this understanding with our analytical results to place bounds on the structural parameters, as discussed above. We explore this approach with an application in Section 5.3.

On the other hand, there are settings in which the econometrician may have reason to think that the primary components of costs are uncorrelated with the drivers of demand. For example, consider demand for coffee beans in the United States. The cost of production primarily depends on weather conditions and other agricultural concerns in Brazil, Colombia, Vietnam, and Guatemala.[14] Demand in the United States is largely unrelated to such factors. Given these considerations, it may be reasonable to assume that demand shocks are orthogonal to marginal costs shocks.

Further, in other cases, the correlation between demand and supply shocks can be accounted for with fixed effects or by explicitly modeling features that drive the correlation. In Section 3.1, we discuss an example with online retail where fixed effects control for the correlation in demand and supply conditions. Similar arguments may be made for consumer products sold in brick-and-mortar stores or electronic goods (Döpper et al., 2023; De los Santos et al., 2021). Later in this paper, we provide an extension to show how a known marginal cost function can be incorporated in the model (Section 4.2). In an application to the cement industry, we find that, after upward-sloping marginal costs are specified in the supply model, the covariance restriction approach yields an estimate that is similar to that obtained when using the instrumental variables employed in the literature (Section 5.2).

We now discuss the requirement that the econometrician specify the supply side of the

[14]Potential supply shocks include: severe frosts, high temperatures, below-average rainfall, excessive rainfall, plant diseases, pests, and fertilizer costs. https://apps.fas.usda.gov/psdonline/circulars/coffee.pdf

model. In practice, empirical research in industrial organization often employs supply models to calculate markups or conduct counterfactuals. For these purposes, jointly imposing the relevant supply model when estimating demand is not a substantive additional assumptions about the economic environment. However, there are cases in which demand-side estimates are of interest independently, and the researcher may be hesitant to imposing a supply model in these cases. To explore misspecification bias, we perform Monte Carlo exercises with a logit demand system and a misspecified supply model, which we detail in Appendix C. If we assume Bertrand competition when the true model features joint profit maximization (i.e., perfect collusion), then the average bias in the price parameter is less than four percent. The potential for bias appears to be mitigated by the fact that the covariance restrictions approach also uses the demand side of the model. These simulations indicate that, in some cases, the bias from supply-side misspecification may not be large.

Finally, it is also worthwhile to consider the assessment of validity in comparison to an instruments-based approach. For the purpose of estimating demand, an instruments-based approach that uses a cost shifter only requires an informal understanding of supply, and thus there is a lower risk that supply-side misspecification could materially affect the estimates. On the other hand, a key similarity between the covariance restrictions approach and the instruments approach is that both require assessments about the distribution of unobservables. In the case of instruments, this assessment manifests as the exclusion restriction. We explore the connection between covariance restrictions and instruments in the next section.

3 Relationship to Instrumental Variables

The covariance restriction approach to estimation interprets observed endogenous variation in quantities and prices through the lens of the economic model. The approach differs from the instrumental variables approach, which seeks to isolate exogenous variation in prices. It also employs different assumptions than those used to obtain "residual instruments" or Hausman instruments. We provide formal distinctions in this section and include Monte Carlo simulations to illustrate.

3.1 Motivating Example

We begin with an economic model for which a traditional instrumental variable approaches and the covariance restriction approach can yield consistent estimates. Consider a simplified setting in which a profit-maximizing monopolist faces linear demand in each of $t = 1, \dots, T$ markets. The demand schedule takes the form:

$$q_{jt} = \alpha p_{jt} + \xi_j + \xi_t + \Delta\xi_{jt} \tag{18}$$

and marginal costs take the form

$$mc_{jt} = \eta_j + \eta_t + \Delta\eta_{jt}, \tag{19}$$

where quantity demanded (q_{jt}) and price (p_{jt}) are observed (we retain the subscript j for notational consistency). Thus, the model includes product- and market-specific shocks (absorbing the constant), but no observable characteristics other than price.

The key identification challenge for estimating α is that prices are often chosen by firms to reflect unobserved demand shocks. In our example, equilibrium prices are given by $p_{jt} = \frac{1}{2}mc_{jt} + \frac{1}{2|\alpha|}\left(\xi_j + \xi_t + \Delta\xi_{jt}\right)$. Thus, prices are higher for higher quality products ($Cov(p_{jt}, \xi_j) > 0$) and in high-demand markets ($Cov(p_{jt}, \xi_t) > 0$). In settings with sufficient observations, these correlations can be accounted for with fixed effects. However, the primary concern about price endogeneity remains as long as $Cov(p_{jt}, \Delta\xi_{jt}) \neq 0$.

A standard approach to resolve price endogeneity is to obtain auxiliary data about a component of costs that is orthogonal to $\Delta\xi_{jt}$. For illustrative purposes, suppose that one could measure $\Delta\eta_{jt}$ directly. Then, $\Delta\eta_{jt}$ would be a valid instrument for p_{jt} in the demand equation as long as $Cov(\Delta\xi_{jt}, \Delta\eta_{jt}) = 0$. It would typically not be sufficient to use a measure of η_j or η_t as an instrument because these components of costs would be absorbed by the fixed effects used to control for product quality and market-level demand differences.[15]

Alternatively, Proposition 1 indicates that one could use the restriction $Cov(\Delta\xi_{jt}, \Delta\eta_{jt}) = 0$ directly in estimation. To provide intuition about the conditions under which this is valid, we build on the economic model using the following stylized example.

Consider an online retailer that sells coffee tables made from two different materials, e.g., wicker and solid wood. Consumers may prefer one product to another (ξ_j) and overall demand for the retailer's products may vary across markets (ξ_t). The online retailer sells the products for different prices in each market, which reflect the above features and also idiosyncratic variation in tastes for products across markets, given by $\Delta\xi_{jt}$. On the supply side, products vary in procurement and distribution costs (η_j), and marginal costs vary across markets due to differences in distribution networks and fuel costs (η_t). For the online retailer, residual product-market variation in costs is due to the interaction of product characteristics with features of the local distribution networks. Specifically, similarly-sized coffee tables can differ significantly in weight, depending on the material. Thus, the product-market cost variation ($\Delta\eta_{jt}$) is approximately given by weight$_j$ $\times$ (fuel cost)$_t$.

[15] In practice, the econometrician often observes only a portion of marginal costs, in which case the instrument z_{jt} can be expressed as a component of the full structural error, $\Delta\eta_{jt} = z_{jt} + \tilde{\Delta\eta}_{jt}$. The unobserved component, $\tilde{\Delta\eta}_{jt}$, may be interpreted as measurement error—for example, if the interaction of fuel costs with weight is only a first-order approximation to actual shipping costs. Isolating a component of costs is an advantage when $Cov(\Delta\xi_{jt}, z_{jt}) = 0$ but $Cov(\Delta\xi_{jt}, \tilde{\Delta\eta}_{jt}) \neq 0$ as the instrument still yields a consistent estimate while the covariance restriction approach may be biased. This bias becomes small when the orthogonal component z explains a greater share of idiosyncratic cost shocks. Thus, it can be helpful to understand the components that contribute the most to the marginal cost residual, even if they are unobserved in data.

Demand for these products does not have any obvious link to idiosyncratic fluctuations in distribution costs. Based on this, one could estimate demand by first obtaining data on product-level characteristics (weight) and market-level features (fuel costs), and then using the interaction of the two to generate cost-shifter instruments in a standard instrumental variables approach. When controlling for product quality and market-level demand, it would be necessary to construct a measure with idiosyncratic across-market variation by product; otherwise, the instrument would be fully absorbed by the fixed effects.

The covariance restrictions approach would leverage the same logic—that idiosyncratic product-market differences in costs are orthogonal to the idiosyncratic product-market differences in preferences—to obtain identification. A key distinction is that the covariance restrictions approach applies the identifying assumption $Cov(\Delta\xi_{jt}, \Delta\eta_{jt}) = 0$ directly, without the need to collect additional data and construct an instrument. This logic is reflected in the approach taken by Döpper et al. (2023) when estimating demand for consumer products. After using fixed effects to account for obvious linkages between demand and costs, the residual supply-side structural error features product-specific changes in input costs and distribution costs, both of which have been exploited as instruments in recent research (e.g., Miller and Weinberg, 2017; Backus et al., 2021). Thus, these examples—along with the applications in Section 5—demonstrate how a similar justification for the validity of instrumental variables may be used to motivate the covariance restrictions approach.

3.2 Excluded Instruments

We now provide formal distinctions between the sets of assumptions that underlie each approach. An instrument is an observable variable that satisfies an exclusion condition and a relevance condition (e.g., Wooldridge, 2010). Using the model of equations (18) and (19), these two conditions can be expressed as:

$$\mathbb{E}[\Delta\xi_{jt} z_{jt}] = 0 \tag{20a}$$

$$\mathbb{E}[p^*_{jt} z_{jt}] \neq 0, \tag{20b}$$

where, again, p^* denotes the residuals from the linear projection of p on $\tilde{x}$. We focus on the case of a single instrument, z. Without loss of generality, we express z_{jt} as a component of the supply-side structural error, $\Delta\eta = z + \tilde{\Delta\eta}$, where $\tilde{\Delta\eta}$ is the remaining unobserved component.[16]

There are two perhaps obvious differences between the instrumental variables approach

[16]The instrument must be linearly independent from $\tilde{x}$, otherwise equation (20b), which can alternatively be expressed as $\mathbb{E}[p^*_{jt} z^*_{jt}]$, would be violated. For the model in equations (18) and (19), $\tilde{x}$ consists (only of) a constant and dummy variables to capture market and product fixed effects. When the model incorporates product-market varying characteristics (x), a variable x^k is a candidate supply-side instrument when $\gamma_k \neq 0$ and $\beta_k = 0$. In this case, we can simply re-define $\tilde{x}$ such that it does not include x^k. This maps z to the above interpretation (as a component of $\Delta\eta$) and avoids the need for more cumbersome notation in this section. Note that omitting x^k from $\tilde{x}$ and including it in $\Delta\eta$ does not affect the validity of a covariance restriction when $z = x^k$ satisfies condition (20a).

to identification and the covariance restrictions approach. The first is that the instrumental variables approach requires that the excluded instrument be observed in the data, as the corresponding estimators such as two-stage least squares (2SLS) are constructed as a function of z. The second is that there may be settings in which the orthogonality condition $\mathbb{E}[\Delta\xi_{jt} z_{jt}] = 0$ holds but the covariance restriction $\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{jt}] = 0$ does not.

There is a third distinction that is more related to the relevance condition of equation (20b). The condition states that some residual variation in prices is explained by the instrument. To show this, the probability limit of the 2SLS estimate of α given the instrument z is:

$$\alpha^{IV} = \frac{Cov(h^*, z)}{Cov(p^*, z)} = \frac{Cov(h^*, \hat{p}^*)}{Var(\hat{p}^*)} \tag{21}$$

where $\hat{p}^*$ is the residual from the linear projection of $\hat{p}$ on $\tilde{x}$, and $\hat{p}$ is defined as the predicted values from the linear projection of p on z and $\tilde{x}$.[17] The first expression shows that the 2SLS estimate equals the ratio of the coefficient obtained in the reduced-form regression of h^* on z to the coefficient obtained in a first-stage regression of p^* on z. The second expression provides a reformulation that is useful for our purposes: the 2SLS estimate equals the coefficient from a regression of h^* on $\hat{p}^*$, the (residualized) first-stage predicted values.

For α^{IV} to be well defined, the relevance condition must hold so that $Cov(p^*, z) \neq 0$. That is, z must add explanatory power for prices above and beyond the variables in $\tilde{x}$. An implication is that choosing the instrument $z_{jt} = \Delta\eta_{jt}$ will not work when there is no residual variation in costs, i.e., $\mathbb{E}[p^*_{jt}\Delta\eta_{jt}] = 0$. In this case, the covariance restrictions approach can still obtain point identification.

Another way to compare the relevance condition to the assumptions needed for the covariance restrictions approach is in terms of the requirement on the denominator of the second expression, $Var(\hat{p}^*) > 0$. Inspection of equation (11) shows that the covariance restriction estimator has a similar-looking requirement, as $Var(p^*)$, which appears in the denominators, must be greater than zero.[18] The condition that $Var(p^*) > 0$, i.e., that there exists residual variation in price after a projection on $\tilde{x}$, is weaker than the relevance condition for instrumental variables. When $Var(\hat{p}^*) > 0$, the second condition $Var(p^*) > 0$, is strictly implied. To see this, note that, by way of counterexample, if $\tilde{x}$ perfectly predicts p such that $Var(p^*) = 0$, there is no residual variation for an instrument to explain.

Moreover, even if condition (20b) is satisfied in the limit, the instrumental variables estimator can exhibit asymptotic bias in finite samples (e.g., Keane and Neal, 2022). This is an important consideration in practice, and many papers have been devoted to address the "weak instrument" problem when this condition is tenuously satisfied (e.g., Bound et al., 1995; Staiger and Stock, 1997; Stock and Yogo, 2005). The covariance restriction approach can side-step this

[17] In terms of data, $\hat{\boldsymbol{P}}^* = \hat{\boldsymbol{P}} - \tilde{\boldsymbol{X}}[\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}}]^{-1}\tilde{\boldsymbol{X}}'\hat{\boldsymbol{P}}$ and $\hat{\boldsymbol{P}} = \tilde{\boldsymbol{Z}}[\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{Z}}]^{-1}\tilde{\boldsymbol{Z}}'\boldsymbol{P}$ for $\tilde{\boldsymbol{Z}} = [\boldsymbol{Z}\ \tilde{\boldsymbol{X}}]$.

[18] As we discuss in Section 2.2, $Var(p^*) > 0$ is an implication of $\mathbb{E}[\boldsymbol{v}_t'\boldsymbol{v}_t]$ having full rank.

issue because all of the residual variation in p is used to construct the estimate. From the Frisch-Waugh-Lovell Theorem, we have $Var(p^*) \geq Var(\hat{p}^*)$, and indeed in applications $Var(\hat{p}^*)$ may be much smaller than $Var(p^*)$ if z is constructed from one of many components of costs.

Thus far, we have framed a valid instrument as satisfying $\mathbb{E}[\Delta\xi_{jt} z_{jt}] = 0$. However, in the context of the model, an alternative is use an instrument that satisfies $\mathbb{E}[\Delta\eta_{jt} z_{jt}] = 0$. Such an instrument could be taken from the demand-side of the model or constructed based on markup shifters, as in Berry et al. (1995). This approach uses the supply side of the model, specifically the first-order conditions of equation (5), to estimate the price parameter. We highlight this possibility because it shows how either demand-side variation or supply-side variation can be used to pin down the price parameter, so long as the appropriate exclusion restriction can be applied. The covariance restriction approach exploits both cost-side variation and demand-side variation implicitly using a single restriction.

3.3 Residual Instruments

We now compare our approach to existing results on the identification of simultaneous equations with covariance restrictions. Wooldridge (2010, p. 258) focuses on the case of two linear equations.[19] Building on our motivating example, we obtain the the following linear demand and supply relationships:

$$q_{jt} = \alpha_1 p_{jt} + \xi_j + \xi_t + \Delta\xi_{jt} \qquad \text{(Demand)} \tag{22a}$$

$$q_{jt} = -\alpha_2 p_{jt} + \alpha_2 \left(\eta_j + \eta_t + \Delta\eta_{jt}\right) \qquad \text{(Supply)} \tag{22b}$$

where, again, the supply relationship comes directly from the monopolist's first-order condition for profit maximization. Because we first consider the general simultaneous equations approach, the price coefficients are allowed to vary across the equations.

The "residual instruments" approach to identification relies on the following identifying moments for some observable z:

$$\mathbb{E}[z_{jt}\Delta\eta_{jt}] = 0 \tag{23a}$$

$$\mathbb{E}[z_{jt}p^*_{jt}] \neq 0 \tag{23b}$$

$$\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{jt}] = 0. \tag{23c}$$

The first two moments are analogous to instrumental variables conditions (20a) and (20b), though, in this case, z is a demand-side instrument and excluded from the supply equation. Here we assume that z is an observed component of the demand-side structural error, $\Delta\xi =$

[19] The discussion in Wooldridge (2010) builds on a substantial literature on covariance restrictions in linear simultaneous equation models (e.g., Hausman and Taylor, 1983; Hausman et al., 1987). More recent research generalizes these results (Matzkin, 2016).

$z + \tilde{\Delta\xi}$. The third moment corresponds to the covariance restriction that $Cov(\Delta\xi, \Delta\eta) = 0$.

These three moments identify the model. Estimation is typically described as proceeding in two steps: in the first step, a standard instrumental variables regression with z as the instrument is used to identify α_2 and the supply equation. The estimated residuals, $\widehat{\Delta\eta}$, can then be used as instruments in the demand equation (22a) to obtain a consistent estimate of α_1. The residuals from the supply equation meet the necessary exclusion restriction in the second step due to the covariance restriction. Thus, in this framework, covariance restrictions have been interpreted as providing excluded instruments.[20] As the above demonstrates, in addition to the covariance restriction, the residual instruments approach requires an additional two conditions about the existence of a valid instrument. In practice, it must also be the case that z and $\widehat{\Delta\eta}$ explain p to a substantial degree, otherwise each step could suffer from the weak instruments problem.

By contrast, our approach to estimation with covariance restrictions recognizes a theoretical connection between the slopes of demand and supply that is implied by the economic model: $\alpha_1 = \alpha_2 = \alpha$. In this case, α is point identified with only one restriction: $Cov(\Delta\xi, \Delta\eta) = 0$. With this approach, there is no need for an excluded instrument (z).[21] This provides a path for identification under a different set of assumptions, while also avoiding the finite-sample challenges of weak instruments.

3.4 Hausman Instruments

A number of articles in industrial organization have relied on prices in related markets as instruments in demand estimation (Gandhi and Nevo, 2021). Typically, the "related markets" refer to distinct geographic areas. In that setting, the price of a product in some market s can be a valid instrument for the price of the same product in market t if marginal costs are correlated across markets (e.g., due to shared production facilities) but demand is not. Such instruments often are referred to as "Hausman instruments" due to their use in Hausman (1996).

Here, we assume that the implementation employs product and market fixed effects.[22] Formally, the conditions needed for validity are that there exist pairs of markets t, s such that:

$$\begin{aligned}
\mathbb{E}[\Delta\xi_{jt}\Delta\xi_{js}] &= 0 && \text{(24a)}\\
\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{js}] &= 0 && \text{(24b)}\\
\mathbb{E}[\Delta\eta_{jt}\Delta\eta_{js}] &\neq 0. && \text{(24c)}
\end{aligned}$$

[20]This interpretation has been influential. For example, McFadden states in lecture notes (dated 1999) 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." See `https://eml.berkeley.edu/~mcfadden/e240b_f01/ch6.pdf`.

[21]However, as discussed earlier, either a demand-side or supply-side instrument would also identify the model.

[22]In practice, researchers sometimes use Hausman instruments that reflect variation across all markets, in which case it is necessary to assume that $\xi_j = 0$.

Condition (24a) states that the demand-side error terms are uncorrelated across markets, condition (24b) states that the demand-side error of one market is uncorrelated with the supply-side error term in another market, and condition (24c) states that supply-side error terms are correlated across markets.

If these conditions are satisfied, then p_{js} is a valid (excluded) instrument for p_{jt} in the demand equation. Thus, analogous to the residual instruments approach in the previous section, this approach leverages assumptions about the correlation structure of unobservables to generate excluded instruments. This approach can suffer from the weak instruments problem, as we discuss in Section 3.2.

Similar to our approach, the Hausman instruments are justified by assumptions about the correlation structure of demand and cost shocks. The set of assumptions are distinct: the three above conditions could be met when $\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{jt}] \neq 0$, or conditions (24a)–(24c) may not be satisfied while $\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{jt}] = 0$. One advantage of the covariance restriction approach is that only a single restriction is required. A second advantage is that it avoids the potential weak instrument problem that Hausman instruments may be subject to in finite samples.

Note that the use of p_{js} as an excluded instrument does not necessarily exploit all of the variation implied by conditions (24a)–(24c). As we discuss in 2.3, it is possible to leverage such orthogonality conditions directly with the method of moments, rather than using them to justify an observed variable (p_{js}) as an instrument. We explore generalizations of our approach with these and other covariance restrictions in Section 4.1.

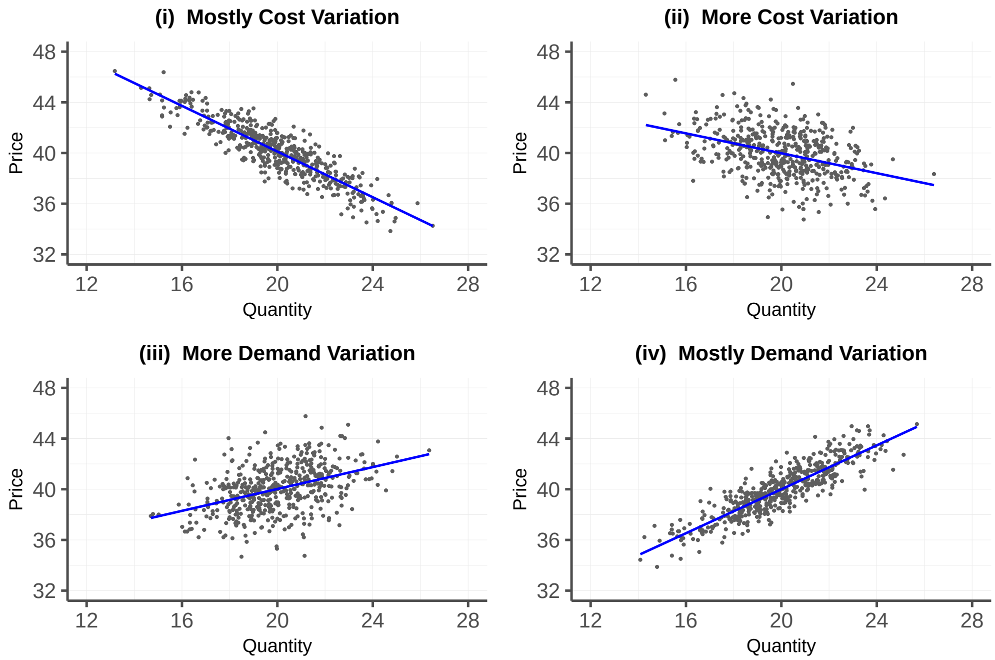
3.5 Finite Sample Comparison

We use Monte Carlo simulations to illustrate the finite sample performance of covariance restrictions relative to excluded instruments. For excluded instruments, we consider both traditional supply-side instruments and also demand-side instruments discussed above in Section 3.2. In both cases, we assume that the available instrument is fully efficient in that it captures all of the relevant exogenous variation—i.e., $z = \Delta\eta$ for a supply-side instrument and $z = \Delta\xi$ for the demand-side instrument. That is, we allow the econometrician to fully observe cost shocks when estimating demand or demand shocks when estimating the supply relationship.

For our simulations, we use the monopoly model of equations (18) and (19) for a single product. We normalize the time fixed effects (ξ_t, η_t) to zero, and we set $\alpha = 1$, $\xi_j = 60$, and $\eta_j = 20$. We assume that $\Delta\xi_{jt}$ and $\Delta\eta_{jt}$ are mean-zero independent normal distributions with standard deviations σ_ξ and σ_η. We consider four specifications: (i) $\sigma_\xi = 1$ and $\sigma_\eta = 4$, (ii) $\sigma_\xi = 2$ and $\sigma_\eta = 3$, (iii) $\sigma_\xi = 3$ and $\sigma_\eta = 2$, and (iv) $\sigma_\xi = 4$ and $\sigma_\eta = 1$. Moving from (i) to (iv), demand-side variation increases and supply-side variation decreases.

As is well known, if both supply and demand variation is present, then equilibrium outcomes provide a "cloud" of data points that need not correspond to the demand curve. To illustrate, we present one simulation of 500 observations from each specification in Figure 1, along with

Figure 1: Prices and Quantities in the Monopoly Model



Notes: This 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.

the fit of an OLS regression of quantity on price. The expected values for the OLS estimator in each scenario are -0.882, -0.385, 0.385, and 0.882. With greater demand-side variation, the endogeneity bias is larger.[23] Using the condition that $Cov(\Delta\xi, \Delta\eta) = 0$, we can correct for endogeneity and construct a covariance restriction estimator for α as $-\sqrt{Var(q)/Var(p)}$ because, in this context, the approximation of equation (16) is exact. The corresponding estimates for each simulation in Figure 1 are -0.98, -1.00, -0.98, and -1.01, close to the true parameter values.

We consider sample sizes of 25, 50, 100, and 500 observations. For each specification and sample size, we randomly draw 10,000 datasets, and with each we estimate the model with a covariance restriction, with a supply-side instrument, and with a demand-side instrument. For the covariance restriction, we estimate α using $-\sqrt{Var(q)/Var(p)}$ as above. For a supply-side instrument, we estimate demand with 2SLS using the cost shock $\Delta\eta$ as the instrument. For the demand-side instrument, we estimate the supply relationship with 2SLS using $\Delta\xi$ as the instrument. All three approaches rely on the same orthogonality condition: $\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{jt}] = 0$.

[23]Inspection of Figure 1 further suggests that there may be connection between OLS bias and goodness-of-fit. Indeed, starting with equation (16), a few lines of additional algebra obtain $\alpha \approx -\left|\alpha^{OLS}\right| / \sqrt{R^2}$ 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	(i) $Var(\eta) \gg Var(\xi)$		(ii) $Var(\eta) > Var(\xi)$		(iii) $Var(\eta) < Var(\xi)$		(iv) $Var(\eta) \ll Var(\xi)$	
25	-1.006	(0.100)	-1.019	(0.198)	-1.017	(0.199)	-1.004	(0.102)
50	-1.003	(0.069)	-1.010	(0.134)	-1.008	(0.136)	-1.002	(0.069)
100	-1.002	(0.047)	-1.005	(0.094)	-1.006	(0.095)	-1.001	(0.049)
500	-1.000	(0.021)	-1.001	(0.041)	-1.001	(0.041)	-1.001	(0.021)

(b) Supply Shifters (IV-1)

Observations	(i) $Var(\eta) \gg Var(\xi)$		(ii) $Var(\eta) > Var(\xi)$		(iii) $Var(\eta) < Var(\xi)$		(iv) $Var(\eta) \ll Var(\xi)$	
25	-1.007	(0.107)	-1.044	(0.314)	-1.273	(3.399)	-0.820	(13.379)
50	-1.003	(0.074)	-1.021	(0.202)	-1.112	(0.623)	-1.369	(10.661)
100	-1.002	(0.050)	-1.010	(0.137)	-1.057	(0.345)	-1.509	(6.676)
500	-1.000	(0.022)	-1.003	(0.060)	-1.009	(0.138)	-1.080	(0.444)

(c) Demand Shifters (IV-2)

Observations	(i) $Var(\eta) \gg Var(\xi)$		(ii) $Var(\eta) > Var(\xi)$		(iii) $Var(\eta) < Var(\xi)$		(iv) $Var(\eta) \ll Var(\xi)$	
25	-0.835	(12.357)	-1.303	(3.667)	-1.040	(0.315)	-1.005	(0.109)
50	-1.299	(11.845)	-1.116	(0.561)	-1.018	(0.203)	-1.003	(0.073)
100	-1.557	(6.517)	-1.052	(0.343)	-1.012	(0.139)	-1.001	(0.052)
500	-1.071	(0.420)	-1.011	(0.137)	-1.002	(0.060)	-1.001	(0.023)

Notes: Results are based on 10,000 simulations of data for each specification and number of observations. The demand curve is $q_{jt} = \alpha p_{jt} + \xi_{jt}$ with $\alpha = -1$ and $\xi_{jt} = \xi_j + \Delta\xi_{jt}$. Marginal costs are $mc_{jt} = \eta_{jt}$ where $\eta_{jt} = \eta_j + \Delta\eta_{jt}$. We consider a single product ($j = 1$) and vary the number of markets/observations from 25 to 500. IV-1 estimates are calculated using 2SLS with cost shocks ($\Delta\eta$) as an instrument in the demand equation. Analogously, IV-2 estimates are calculated using 2SLS with demand shocks ($\Delta\xi$) as an instrument in the supply relationship. We specify $\Delta\xi$ and $\Delta\eta$ as mean-zero independent normal distributions with standard deviations σ_ξ and σ_η. We consider four specifications: (i) $\sigma_\xi = 1$ and $\sigma_\eta = 4$, (ii) $\sigma_\xi = 2$ and $\sigma_\eta = 3$, (iii) $\sigma_\xi = 3$ and $\sigma_\eta = 2$, and (iv) $\sigma_\xi = 4$ and $\sigma_\eta = 1$. Moving from (i) to (iv), demand-side variation increases and supply-side variation decreases.

Table 1 provides the mean and (empirical) standard error of the point estimates for each specification and approach.[24] Panel (a) shows that the covariance restriction approach to estimation yields estimates that are consistently close to the true value. Panel (b) shows that, with supply-side instruments, small-sample bias becomes substantial with smaller datasets and less variance in the cost shock. This is due to a weak instrument—for example, the mean first-stage F-statistics in specification (iv) are 2.8, 4.2, 7.4, and 32.1 for markets with 25, 50, 100, and 500 observations, respectively. Panel (c) shows that, with demand-side instruments, small-sample bias becomes substantial with smaller datasets and less variance in the demand shock, which is also due to a weak instruments problem.

[24]To avoid outliers arising from the weak instrument problem, we bound the estimates of α on the range $[-100, 100]$. For specifications that suffer from weak instruments, this biases the standard errors toward zero. This affects specifications where the estimated standard error is greater than one, i.e., in 8 of 48 specifications.

Thus, in settings where instruments perform poorly, a covariance restriction may still provide a precise estimate when the assumptions about the environment are correct. In our simulations, the covariance restriction has smaller standard deviations than either instrumental variables strategy. Intuitively, even though exogenous variation is not observed, the covariance restrictions approach exploits how variation from demand and supply is reflected in equilibrium prices and quantities.

4 Extensions

In this section, we provide two extensions of our main results. First, we discuss a broader class of potential covariance restrictions. Second, we consider the case of marginal cost functions. We use both extensions in the applications of Section 5.

4.1 Generalized Covariance Restrictions

Thus far, our analysis has focused on covariance restrictions between own demand and cost shocks. Our results demonstrate that this moment is expected to generate a consistent estimate of the price parameter. We now consider different covariance restrictions that generalize the approach. Though these other covariance restrictions do not provide a similar guarantee of point identification, they may work well in certain settings. Additionally, they may pin down other parameters of interest (e.g., those in θ_2), in addition to the price parameter.

Consider the assumptions (24a)–(24c) that are required for the Hausman instruments. Rather than using these assumptions to motivate the use of an instrument, the assumptions could be employed directly in a method-of-moments estimator, where, as in equation (14), the estimated residuals are generated from econometric model for a candidate parameter. This approach has the advantage over the Hausman instruments approach in that the estimator would utilize all of the variation implied by the identifying moments.

Alternatively, 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 $\mathbb{E}_{jt}[\Delta\xi_{jt}^2\Delta\eta_{jt}]$ and $\mathbb{E}_{jt}[\Delta\xi_{jt}\Delta\eta_{jt}^2]$. Or it may be reasonable to assume that average shocks are uncorrelated across groups of products, i.e., $\mathbb{E}_{gt}[\overline{\Delta\xi}_{gt}\overline{\Delta\eta}_{gt}] = 0$, where $\overline{\Delta\xi}_{gt}$ and $\overline{\Delta\eta}_{gt}$ are the mean demand and cost shocks for products in group g.

Finally, it may be useful to consider cross-product covariance restrictions, i.e.,

$$\mathbb{E}_t[\Delta\xi_{jt}\Delta\eta_{kt}] = 0 \quad \forall j \neq k. \tag{25}$$

These restrictions state that the demand shock for product j is uncorrelated with the cost shock for product k. The expectation in equation (25) can be taken over t and k to obtain J restrictions or over markets (as written above) to obtain $J \times (J-1)$ restrictions.[25]

We also note that our analytical results do not require that the value of the covariance terms are equal to zero. Proposition 1 can be used to construct consistent estimates for any ς for which $\mathbb{E}[\Delta\xi_{jt}\Delta\eta_{jt}] = \varsigma$. In certain situations, it may be possible to employ an estimate of the correlation in demand and cost shocks to identify the price parameter. For example, Berry et al. (1995) report that this correlation is 0.17. For a similar empirical setting, it may be reasonable to invoke Proposition 1 to obtain an estimate of α conditional on this value.

4.2 Marginal Cost Functions

If marginal cost depends on quantity then estimation that exploits the covariance restriction $Cov(\Delta\xi, \Delta\eta) = 0$ may not be credible unless the econometrician models the relationship between quantity and marginal cost explicitly. Consider the case in which marginal costs can be expressed as the following function:

$$mc_{jt} = x_{jt}\gamma + g(q_{jt}, \boldsymbol{w}_t; \tau) + \eta_{jt} \tag{26}$$

where $g(\cdot)$ is a function that depends on quantity, the data in $\boldsymbol{w}_t$, and the vector of parameters τ that is contained in θ_2. The supply relationship becomes:

$$\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) = -\alpha p_{jt} + \alpha x_{jt}\gamma + \alpha g(q_{jt}, \boldsymbol{w}_t; \tau) + \alpha\eta_{jt} \tag{27}$$

In this augmented model, both markup adjustments and varying marginal costs contribute to price endogeneity. If the econometrician omits $g(\cdot)$ from the model, then the residual cost shock is $\widetilde{\Delta\eta}_{jt} \equiv g(q_{jt}, \boldsymbol{w}_t; \tau) + \Delta\eta_{jt}$. Then, $Cov(\Delta\xi, \Delta\eta) = 0$ does not imply that $Cov(\Delta\xi, \widetilde{\Delta\eta}) = 0$, and the approach of Section 2 may not produce consistent estimates.

However, tracing through the steps developed in Section 2.2, we can show that α is identified by the covariance restriction $Cov(\Delta\xi, \Delta\eta) = 0$ for any value of τ, given knowledge of $g(\cdot)$. In that scenario, $g(\cdot)$ can be calculated from the data given τ. Let g_{jt} (or simply g) denote the values of $g(\cdot)$ for each j and t, given τ. The OLS regression of h on p and $\tilde{x}$ yields a price coefficient with the following probability limit:

$$\alpha^{OLS} = \alpha - \frac{1}{\alpha}\frac{Cov\left(\Delta\xi, \lambda\right)}{Var(p^*)} + \frac{Cov(\Delta\xi, g)}{Var(p^*)} + \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} \tag{28}$$

This equation can be reformulated such that the demand-side error term, $\Delta\xi$, is replaced with the probability limit of OLS residuals, $\Delta\xi^{OLS}$, creating an analog to equation (10). Rearranging terms and assuming $Cov(\Delta\xi, \Delta\eta) = 0$ then yields an analog to equation (11):

Corollary 1. *If marginal costs take the semi-linear form of equation* (26) *and* $Cov(\Delta\xi, \Delta\eta) = 0$*,*

[25] Similarly, the own-product restrictions may assumed to hold separately by product, providing J restrictions.

then α solves the following quadratic equation:

$$
\begin{aligned}
0 = \quad & \left(1 - \frac{Cov(p^*, g)}{Var(p^*)}\right)\alpha^2 \\
& + \left(\frac{Cov\,(p^*, \lambda)}{Var(p^*)} - \alpha^{OLS} + \alpha^{OLS}\frac{Cov(p^*, g)}{Var(p^*)} + \frac{Cov(\Delta\xi^{OLS}, g)}{Var(p^*)}\right)\alpha \\
& + \left(-\alpha^{OLS}\frac{Cov\,(p^*, \lambda)}{Var(p^*)} - \frac{Cov\left(\Delta\xi^{OLS}, \lambda\right)}{Var(p^*)}\right)
\end{aligned}
$$

Given θ_2, and thus h and g, there are at most two solutions for α, and the lower root can fairly be targeted in most applications. In estimation, the method-of-moments can be used to jointly estimate α and τ, using the covariance restriction to identify α and auxiliary moments to identify τ. The auxiliary moments can consist of excluded instruments or the generalized covariance restrictions discussed above. Thus, it is possible to control for a direct relationship between quantity and marginal costs with additional structure. We explore a cost function approach to estimation in an application to cement (Section 5.2).

5 Empirical Applications

We provide three empirical applications to demonstrate how covariance restrictions can inform inference. The three settings—ready-to-eat (RTE) cereals, cement, and airlines—differ in a variety of ways that influence our implementation. With RTE cereals we proceed with estimation under $Cov(\Delta\xi, \Delta\eta) = 0$, assuming constant marginal costs and using fixed effects to absorb potentially confounding variation, as discussed in Section 3. With cement, capacity constraints imply that marginal costs can increase with quantities. We follow an approach developed in the literature and model this effect explicitly, after which $Cov(\Delta\xi, \Delta\eta) = 0$ becomes credible (as in Section 4.2). Finally, with airlines, the relationship between demand shocks and prices can be complicated; instead of modeling it directly we apply a bounds approach (as in Section 2.4).

5.1 Ready-to-Eat (RTE) Cereals

We choose RTE cereals for our first application because, with panel data and appropriate fixed effects, a covariance assumption appears credible, for reasons that we explain below. Furthermore, it allows us to develop the covariance restrictions approach to estimation in the context of the random coefficients logit demand model (Berry et al., 1995). We use the pseudo-real cereals data of Nevo (2000) and compare estimates obtained with a covariance restriction to those obtained with the provided instruments. There are 24 products, 47 cities, and 2 quarters.[26]

[26] 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

Let the indirect utility that consumer i receives from product j in market t (a combination of a quarter and a city) be given by

$$u_{ijt} = \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \phi_{ijt}(x_{jt}, p_{jt}, \nu_i, D_i; \theta_2) + \epsilon_{ijt} \tag{29}$$

where δ_{jt} denotes a common component and ϕ_{ijt} provides consumer-specific utility as a function of data and θ_2. These components are specified as

$$\begin{aligned}\delta_{jt} &= \alpha p_{jt} + x_{jt}\beta + \xi_j + \Delta\xi_{jt} \\ \phi_{ijt} &= [p_{jt}\ x_{jt}](\Pi D_i + \Sigma \nu_i)\end{aligned}$$

such that consumer-specific utility is linear in the parameters $[\alpha_i; \beta_i] = [\alpha; \beta] + \Pi D_i + \Sigma \nu_i$. Consumers can pick any one of the inside goods ($j = 1, \ldots, 24$) or an outside good ($j = 0$) that provides indirect utility of $u_{i0t} = \epsilon_{i0t}$. Mapping to the notation of Section 2, α and β are contained in θ_1, Π and Σ are in θ_2, and the data $\boldsymbol{x}_t$ and $\boldsymbol{p}_t$ are included in $\boldsymbol{w}_t$.

Demand is expressed in terms of market shares.[27] Let N_t denote the number of consumers in a market, which we assume to be large. Under the assumption that unobserved shock ϵ_{ijt} is distributed i.i.d. type 1 extreme value, the market share for product j in market t ($j \neq 0$) can be written as

$$s_{jt} = \varsigma_{jt}(\boldsymbol{\delta}_t, \boldsymbol{w}_t; \theta_2) \equiv \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\exp(\delta_{jt} + \phi_{ijt}(\boldsymbol{w}_t; \theta_2))}{1 + \sum_{k=1}^{J} \exp(\delta_{kt} + \phi_{ikt}(\boldsymbol{w}_t; \theta_2))} \tag{30}$$

Stacking across products, we obtain the vector-valued equation $\boldsymbol{s}_t = \boldsymbol{\varsigma}_t(\boldsymbol{\delta}_t, \boldsymbol{w}_t; \theta_2)$. Because each $\varsigma_{jt}(\cdot)$ is strictly increasing in δ_{jt}, this equation can be inverted to obtain $\boldsymbol{\delta}_t(\boldsymbol{s}_t, \boldsymbol{w}_t; \theta_2)$. Each element δ_{jt} corresponds to $h_{jt} = h^{(j,t)}(\boldsymbol{s}_t, \boldsymbol{w}_t; \theta_2)$ in the notation of Section 2. Thus, in implementation, the contraction mapping of Berry et al. (1995) can obtain the $J \times 1$ vector $\boldsymbol{h}_t$ given $\boldsymbol{s}_t$, $\boldsymbol{w}_t$, and θ_2.

On the supply side of the model, marginal costs are given by

$$mc_{jmt} = \eta_j + \Delta\eta_{jt} \tag{31}$$

Prices are set by multi-product firms in Bertrand competition. Following the general results for multi-product firms in Appendix A.5, equilibrium markups take the form specified in equation (3), such that $\lambda^{(j,t)}$ can be expressed as a function of $\boldsymbol{s}_t$, $\boldsymbol{w}_t$, and θ_2. For additional details, see Appendix A.2. Note that we follow Nevo (2000) and exclude market fixed effects from demand and supply.

imposing that the parameter on $Price \times Income^2$ takes a value of zero. We refer readers to Nevo (2000) for details on the data.

[27] Following the discussion in Section 2.1, market shares can be converted into quantities using market size: $q_{jt} = s_{jt}M_t$. Thus, if the data have quantities rather then market shares then M_t must be in $\boldsymbol{w}_t$.

We use the covariance restriction $Cov(\Delta\xi_{jt}, \Delta\eta_{jt}) = 0$ in estimation. The supply-side structural error term incorporates some of the cost-shifter instruments that have been used in the recent literature, including time-varying, product-specific shipping costs (Miller and Weinberg, 2017) and the time-varying prices of product-specific ingredients (Backus et al., 2021). Given the fixed effects, these cost shifters can be conceptualized as providing the variation that is exploited in estimation. Furthermore, it may be reasonable to think that marginal costs are roughly constant with consumer products, as often maintained in the literature (Villas-Boas, 2007; Chevalier et al., 2003; Hendel and Nevo, 2013; Miller and Weinberg, 2017; Backus et al., 2021).

The parameters for estimation include α and β (contained in θ_1) and also Π and Σ (contained in θ_2). Identification of θ_1 is obtained under the exogeneity of $\tilde{x}$ as conditions (1) and (5) are satisfied. Additional identifying assumptions are needed for θ_2. Some recent applications use micro-moments constructed from the observed behavior of individual consumers (e.g., Backus et al., 2021; Döpper et al., 2023) or "second-choice" data on what consumers view as their next-best option (e.g., Grieco et al., 2023). Both of these strategies identify Π and Σ but not the price parameter (Berry and Haile, 2020). This separability allows for a two-step approach to estimation in which the price parameter is estimated after the other parameters. An alternative strategy is to use instruments constructed from competitor characteristics (e.g., Berry et al., 1995; Gandhi and Houde, 2020) to identify the additional parameters. As none of these options are available to us given the data and specification, we pursue an alternative approach based on a generalization of the covariance restriction assumption.

Specifically, we extend the assumption that residual demand and cost shocks are uncorrelated to all cross-product pairs, such that $Cov(\Delta\xi_{jt}, \Delta\eta_{kt}) = 0$ for all j, k. The joint restrictions are valid if the demand shock of each product is orthogonal to its own marginal cost shock and those of all other products. As there are 24 products in each market, the full covariance matrix of demand and cost shocks provides sufficient moments to estimate the 12 nonlinear parameters in the specification.

Table 2 summarizes the results of estimation based on the instruments (panel (a)) and covariance restrictions (panel (b)). In the application, D_i consists of measures of log income, age, and an indicator for whether the individual is a child, ν_i is drawn as a standard normal to capture unobserved demographics, and x contains a constant, sugar content, and an indicator for whether the cereal gets mushy with milk. 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. The standard errors are noticeably smaller with covariance restrictions, which likely reflects that the covariance restrictions approach to estimation more fully exploits the variation that is present in the data.

Table 2: Point Estimates for Ready-to-Eat Cereal

(a) Available Instruments

		Standard	Interactions with Demographics		
Variable	Means	Deviations	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

		Standard	Interactions with Demographics		
Variable	Means	Deviations	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: This table reports point estimates for the random-coefficients logit demand system estimated using the Nevo (2000) dataset. Panel (a) employs the available instruments. Panel (b) employs covariance restrictions.

We conclude that in this setting—where a covariance restriction appears credible—estimation with covariance restrictions and with instruments indeed produce similar results.

5.2 The Portland Cement Industry

Our second empirical application considers a setting in which marginal costs increase with output. We build on the marginal cost specification from Section 4.2, in which the upward-sloping part of the cost function can be modeled explicitly. To illustrate, we consider the setting and data of Fowlie et al. (2016) ["FRR"], which examines market power in the cement industry.

The data are a balanced panel of 520 region-year observations for 20 regions over 1984-2009, with the regions corresponding to selected urban areas in the United States. There are an average of 4.65 cement firms located in each region-year.[28]

The model features Cournot competition among cement plants facing capacity constraints.

[28] See FRR for details on the data, which are available for downloaded as part of the replication package.

Let the market demand curve in region r and year t have a logit form:

$$h^{(rt)}(Q_{rt}, \boldsymbol{w}_{rt}; \theta_2) \equiv \ln(Q_{rt}) - \ln(M_r - Q_{rt}) = x_{rt}\beta + \alpha p_{rt} + \xi_r + \Delta\xi_{rt} \tag{32}$$

where $Q_{rt} = \sum_{j\in\mathcal{J}} q_{jrt}$ is total quantity produced in the region-year, M_r is the "market size" of the region (and is contained in θ_2), and the only characteristic in x_{rt} is a constant.[29] Further, we allow marginal costs to vary with quantity according to

$$mc_{jrt} = x_{rt}\gamma + \alpha p_{rt} + g_{jrt}(q_{jrt}, \boldsymbol{w}_{rt}; \tau) + \Delta\eta_{jrt} \tag{33}$$

We follow FRR in the specification of the cost function and included fixed effects. In particular, we assume that that g is a "hockey stick" function, $g_{jrt}(q_{jrt}, \boldsymbol{w}_{rt}; \tau) \equiv 2\tau 1\{q_{jrt}/k_{jr} > 0.9\}(q_{jrt}/k_{jr} - 0.9)$, where k_{jr} and q_{jrt}/k_{jr} are capacity and utilization, respectively. Marginal costs are constant if utilization is less than 90%. Above this threshold, marginal costs increase linearly in quantities at a rate determined by $\tau \geq 0$. Mapping to the notation of Section 2, τ is a scalar element of θ_2, and $\boldsymbol{w}_{rt}$ includes k_{jr} for each firm (in addition to M_r).

As in our baseline model, correlation between price and the demand-side structural error term can arise both due to markup adjustments and the effect of demand on marginal costs. However, due to the presence of $g_{jrt}(\cdot)$ in the cost function, the latter channel exists even under the covariance restriction $Cov(\Delta\xi_{rt}, \overline{\Delta\eta}_{rt}) = 0$, where $\overline{\Delta\eta}_{rt} = \frac{1}{J}\Delta\eta_{jrt}$. If $g_{jrt}(\cdot)$ is known or can identified with additional moments, then the covariance restriction is sufficient to resolve price endogeneity, as the model informs the markup adjustments. In estimation, we maintain the covariance restriction at the market level.

Our demand and supply framework of equations (1) and (5) readily admits Cournot competition. As only market-level price and costs measures are observed, one must use the mean firm-level quantity $\overline{q}_{rt} = \frac{1}{J}Q_{rt}$ to obtain an expression for mean market-level markups and λ. In particular, when firms compete in quantities, we obtain $\lambda_{rt} = \frac{1}{J}\frac{dh}{dq}Q_{rt}$. Section 4.2 establishes the necessary results to incorporate increasing marginal costs in our framework. In our implementation, we assume that $\psi = 800$, such that our $g_{jrt}(\cdot)$ function is close to what is used in Fowlie et al. (2016), and then use a method-of-moments estimator.

In the context of the cement industry, whether the covariance restriction is reasonable may depend primarily on the relationship between construction activity (a shifter of unobserved demand) and the prices of coal and electricity (determinants of unobserved marginal cost). There is a theoretical basis to assume orthogonality: for example, if coal suppliers have limited market power and roughly constant marginal costs, then coal prices should not respond much to demand for cement. Indeed, this is the identification argument of FRR, as coal and electricity prices are included in the set of excluded instruments. Consistent with this, the time-series of

[29] We use logit demand rather than the constant elasticity demand of FRR to allow for adjustable markups. The 2SLS results are unaffected by the choice. In our implementation, we assume $M_r = 2 \times \max_t\{Q_{rt}\}$.

coal prices over 1980-2010 is not obviously correlated with macroeconomic conditions (e.g., Miller et al., 2017).

We find that the covariance restrictions approach yields a demand elasticity of -1.15, with a standard error of 0.18.[30] This is close to the 2SLS estimate of -1.07 (standard error 0.19) that we obtain using the FRR instruments: coal prices, natural gas prices, electricity prices, and wage rates. That the two approaches generate similar estimates may reflect that the identifying assumptions themselves are similar, with the main difference being whether the cost shifters are treated as observed (2SLS) or unobserved (covariance restrictions). By contrast, we obtain a demand elasticity of -0.47 and a standard error of 0.15 using OLS. If we use the covariance restriction approach without accounting for the presence of $g_{jrt}(\cdot)$, we obtain a demand elasticity of -0.90 and a standard error of 0.13, which is in between the OLS and 2SLS estimates and demonstrates how accounting for marginal cost functions can matter for estimation results.

5.3 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"]. The economics of the industry suggest that the covariance restriction $Cov(\Delta\xi, \Delta\eta) = 0$ would not be credible. The reason is that airlines bear an opportunity cost when they sell a seat because it can no longer be sold at a higher price to another passenger (Williams, 2022). Thus, all else equal, greater demand generates higher marginal costs, inclusive of the opportunity cost. Absent a model of these opportunity costs, would be difficult to achieve point identification using the covariance restriction. Instead, we illustrate how to proceed in such cases by constructing bounds on the parameters of interest.

The data feature 2,950 city-pairs in the United States observed over the four quarters of 2004. A market is a city-pair-quarter combination. Products are classified into the following groups: nonstop flights, one-stop flights, and the outside good. On average, there are 7.98 products per market (not including the outside good) including 2.04 nonstop products.[31]

The nested logit demand system can be expressed as

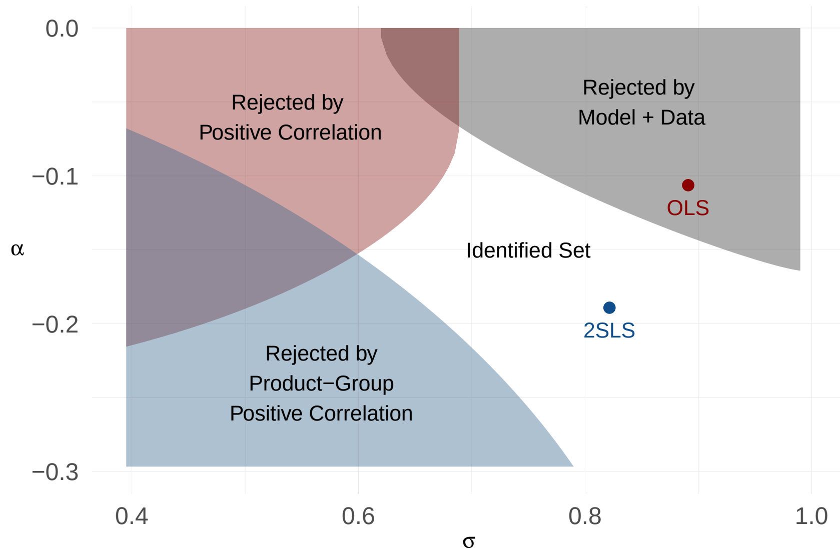
$$h^{(j,mt)}(\boldsymbol{q}_{mt}, \boldsymbol{w}_{mt}; \theta_2) \equiv \ln s_{jmt} - \ln s_{0mt} - \sigma \ln \overline{s}_{jmt|g} = \alpha p_{jmt} + x_{jmt}\beta + \xi_{a(j)} + \xi_{mt} + \Delta\xi_{jmt} \quad (34)$$

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 $j = 0$. Consumer preferences vary by airline ($\xi_{a(j)}$) and by route-quarter (ξ_{mt}), which we account for with fixed effects. Higher

[30] We obtain bootstrapped standard errors based on 200 random samples constructed by drawing from the data with replacement.

[31] We thank Victor Aguirregabiria for providing the data, which derive from the DB1B database maintained by the Department of Transportation. 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.

Figure 2: Analysis of Bounds in the Airlines Industry



Notes: This figure displays candidate parameter values for (σ, α). 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 $Cov(\Delta\xi, \Delta\eta) < 0$, and the blue region indicates parameters that generate $Cov(\overline{\Delta\xi}, \overline{\Delta\eta}) < 0$. 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 $[0, 1)$.

values of σ increase within-group consumer substitution relative to across-group substitution.[32]

We impose three sets of bounds. First, we assume that product-level demand and cost shocks are weakly positive, i.e., $Cov(\Delta\xi_{jmt}, \Delta\eta_{jmt}) \geq 0$, based on the role of opportunity costs in the industry. Second, if the correlation in product-level shocks is weakly positive, it is reasonable to also assume that the correlation in group-level shocks is also weakly positive. That is, overall demand for non-stop flights in a market may drive up the opportunity costs for non-stop flights. Thus, building on Section 4.1, we apply the group-level inequality

$$E_{gmt}[\overline{\Delta\xi}_{gmt}\overline{\Delta\eta}_{gmt}] \geq 0, \tag{35}$$

where $\overline{\Delta\xi}_{gmt} = \frac{1}{|g|}\sum_{j\in g}\Delta\xi_{jmt}$ and $\overline{\Delta\eta}_{gmt} = \frac{1}{|g|}\sum_{j\in g}\Delta\eta_{jmt}$ are the mean demand and cost shocks within a group-market-period. Finally, we combine these bounds with the model-based bounds developed in Section 2.4. We then construct an identified set by rejecting values of the parameters (α, σ) that fail to generate the data or that deliver negative correlations between costs and demand.

[32]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. In estimation, we include airline fixed effects and route×quarter fixed effects. 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 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 α if $\sigma \geq 0.62$. As σ becomes larger, a more negative α 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 $Cov(\Delta\xi_{jmt}, \Delta\eta_{jmt}) \geq 0$. The dark blue region provides the corresponding set for the prior $Cov(\overline{\Delta\xi}_{gmt}, \overline{\Delta\eta}_{gmt}) \geq 0$ 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 α, as these lower values cannot generate positive correlation in both product-level and product-group-level shocks. Thus, the bounds serve to reject the logit model ($\sigma = 0$) in favor of nested logit, even with relatively limited information about the covariance of shocks.

Similarly, we obtain an upper bound on α of -0.067 across all values of σ. Combined, these bounds indicate that the mean own-price elasticity is less than -0.537. 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.

6 Conclusion

We have shown that covariance restrictions between unobserved demand and cost shocks can resolve price endogeneity and allow for consistent estimation in models of imperfect competition. The covariance restrictions approach is notable in part because, unlike approaches with instrumental variables, an exogenous portion of price is not isolated and then used in estimation; 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.

More broadly, our analysis shows how imposing a supply-side model provides feasible paths to identification. In addition to covariance restrictions, our formal results illustrate how demand-side instruments can be sufficient to resolve price endogeneity. We also establish model-free bounds, in which the model and the data jointly can reject certain values of the price parameter, without the need for additional identifying assumptions. Conditional on meeting these bounds, there is typically a unique mapping between the price coefficient and the covariance of demand and costs shocks. In such settings, the covariance term can act as a free parameter to rationalize different values of the price coefficient.

The appeal of the covariance restrictions approach relative to alternatives will depend on data availability and the institutional details of the industry under study. In cases where cost shifters are observed, instrumental variables has the advantage that only an informal understanding of supply is required to estimate demand. By contrast, the covariance restrictions approach leverages all of the observed variation in prices and quantities, but it requires a formal supply-side model. Our results provide paths to identification that may facilitate research in areas for which strong supply-side instruments are unavailable. As is true with most empirical work, the reliability of research that employs a particular approach 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, 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, 1749–1797.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," Econometrica, 63, 841.

BERRY, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," RAND Journal of Economics, 25, 242–262.

BERRY, S. T. AND P. A. HAILE (2020): "Nonparametric Identification of Differentiated Products Demand Using Micro Data," Working paper.

BOUND, J., D. A. JAEGER, AND R. A. BAKER (1995): "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak," Journal of the American Statistical Association.

BRESNAHAN, T. F. (1982): "The Oligopoly Solution Concept is Identified," Economics Letters, 10, 87–92.

BRODA, C. AND D. E. WEINSTEIN (2006): "Globalization and the Gains from Variety," The Quarterly Journal of Economics, 121, 541–585.

——— (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.

CHEVALIER, J. A., A. K. KASHYAP, AND P. E. ROSSI (2003): "Why Don't Prices Rise during Periods of Peak Demand? Evidence from Scanner Data," American Economic Review, 93, 15–37.

CHO, S., G. LEE, J. RUST, AND M. YU (2018): "Optimal Dynamic Hotel Pricing," Working paper.

CONLON, C. AND J. GORTMAKER (2020): "Best Practices for Differentiated Products Demand Estimation with PyBLP," RAND Journal of Economics, 51, 1108–1161.

De Loecker, J. (2011): "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," Econometrica, 79, 1407–1451.

De los Santos, B., D. P. O'Brien, and M. R. Wildenbeest (2021): "Agency Pricing and Bargaining: Evidence from the E-Book Market," Working Paper.

Deaton, A. and J. Muellbauer (1980): "An Almost Ideal Demand System," The American Economic Review, 70, pp. 312–326.

Dixit, A. K. and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," American Economic Review, 67, 297–308.

Döpper, H., A. MacKay, N. Miller, and J. Stiebale (2023): "Rising Markups and Consumer Preferences," Working paper.

Doraszelski, U. and J. Jaumandreeu (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, 2231–2267.

Feenstra, R. C. (1994): "New Product Varieties and the Measurement of International Prices," American Economic Review, 157–177.

Fisher, F. (1963): "Uncorrelated Disturbances and Identifiability Criteria," International Economic Review, 4, 134–152.

——— (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, 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," Working paper.

Gandhi, A. and A. Nevo (2021): "Empirical Models of Demand and Supply in Differentiated Products Industries," In Ho, K., Hortacsu, A., and Lizzeri, A. (Eds.), *Handbook of Industrial Organization, 4th Edition*. Elsevier.

GRIECO, P. L. E., C. MURRY, AND A. YURUKOGLU (2023): "The Evolution of Market Power in the US Automobile Industry," Quarterly Journal of Economics, forthcoming.

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, 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, 1527–1549.

HENDEL, I. AND A. NEVO (2013): "Intertemporal Price Discrimination in Storable Goods Markets," American Economic Review, 103, 2722–2751.

KEANE, M. AND T. NEAL (2022): "A Practical Guide to Weak Instruments," Working Paper.

KNITTEL, C. R. AND K. METAXOGLOU (2014): "Estimation of Random-Coefficients Demand Models: Two Empiricists' Perspective," The Review of Economics and Statistics, 96, 34–59.

KOOPMANS, T. C. (1949): "Identification Problems in Economic Model Construction," Econometrica, 17, 125–144.

KOOPMANS, T. C., H. RUBIN, AND R. B. LEIPNIK (1950): Statistical Inference in Dynamic Economic Models (Cowles Commission Monograph 10), New York: John Wiley & Sons, chap. Measuring the Equation Systems of Dynamic Economics.

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, 319–327.

LEONTIEF, W. (1929): "Ein Versuch Zur Statistichen Analyse von Angebot und Nachfrage," Weltvirtschaftliches Archiv, 30, 1–53.

LI, S., J. MAZUR, Y. PARK, J. ROBERTS, A. SWEETING, AND J. ZHANG (2022): "Repositioning and Market Power after Airline Mergers," RAND Journal of Economics, 53, 166–199.

Matzkin, R. L. (2016): "On Independence Conditions in Nonseparable Models: Observable and Unobservable Instruments," Journal of Econometrics, 191, 302–311.

Miller, N., M. Osborne, and G. Sheu (2017): "Pass-Through in a Concentrated Industry: Empirical Evidence and Regulatory Implications," RAND Journal of Economics, 48, 69–93.

Miller, N. H. and M. Weinberg (2017): "Understanding the Prices Effects of the MillerCoors Joint Venture," Econometrica, 85, 1763–1791.

Nevo, A. (1998): "Identification of the Oligopoly Solution Concept in a Differentiated-Products Industry," Economics Letters, 59, 391–395.

——— (2000): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," Journal of Economics & Management Strategy, 9, 513–548.

——— (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," Econometrica, 69, 307–342.

Nevo, A. and A. Rosen (2012): "Identification with Imperfect Instruments," Review of Economics and Statistics, 94, 659–671.

Petterson, M. S., D. G. Seim, and J. M. Shapiro (2022): "Bounds on a Slope from Size Restrictions on Economic Shocks," American Economic Journal: Microeconomics, forthcoming.

Romeo, C. (2010): "Filling Out the Instrument Set in Mixed Logit Demand Systems for Aggregate Data," Working Paper.

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, 1–17.

Staiger, D. and J. H. Stock (1997): "Instrumental Variables with Weak Instruments," Econometrica, 65, 557–586.

Stock, J. H. and M. Yogo (2005): "Testing for Weak Instruments in Linear IV Regression," *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*.

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, 625–652.

WALDFOGEL, J. (2003): "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets," RAND Journal of Economics, 34, 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. (2022): "The Welfare Effects of Dynamic Pricing: Evidence from Airline Markets," Econometrica, 92, 831–858, forthcoming.

WOOLDRIDGE, J. M. (2010): *Econometric Analysis of Cross Sectional and Panel Data*, Second Edition, MIT Press.

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, 763–771.

Appendix

A Demand System Applications

The demand system of equation (1) is sufficiently flexible to nest monopolistic competition with linear demand (e.g., as in the motivating example) and the discrete choice demand models that support much of the empirical research in industrial organization. The demand assumption can also be modified to allow for semi-linearity in a transformation of prices, $f(p_{jt})$:

$$h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) = \alpha f(p_{jt}) + x_{jt}\beta + \xi_{jt} \tag{A.1}$$

Under this modified assumption, it is possible to employ a method-of-moments approach to estimate the structural parameters. In certain cases, it straightforward to extend our analytical results.

For example, when $f(p_{jt}) = \ln p_{jt}$, we can obtain our identification results under a modified assumption about the structure of costs. The optimal price in such demand systems takes the multiplicative form $p_{jt} = \mu_{jt} \cdot mc_{jt}$, where μ_{jt} is a markup that reflects demand parameters and (possibly) demand shocks. Assume that log marginal costs are linear in characteristics, such that $\ln mc_{jt} = x_{jt}\gamma + \eta_{jt}$. As in Section 2.2, consider the augmented exogenous characteristics $\tilde{x}$ to include a full set of dummy variables for products and markets. The probability limit ($T \to \infty$) of the OLS estimate of α obtained from a regression of h on $\ln p$ and $\tilde{x}$ is given by:

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

This expression is analogous to equation (D.1). Therefore, the results developed in this paper extend in a straightforward manner.

We provide some typical examples below for single-product firms with Bertrand competition. We then show how multi-product firms and other models of competition fit within the framework of Section 2.

A.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

$$h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) \equiv \ln(s_{jt}) - \ln(s_{0t}) - \sigma \ln(\overline{s}_{j|g,t})$$

where $\overline{s}_{j|g,t} = \sum_{j \in \mathscr{J}_g} \frac{s_{jt}}{\sum_{j \in \mathscr{J}_g} s_{jt}}$ is the market share of firm j within its group. Shares are obtained by dividing quantities by the market size M_t. The market size and group identities are contained in $\boldsymbol{w}_t$. The parameter $\sigma \in [0, 1)$ is the only element of θ_2, and it determines the extent to which consumers substitute disproportionately among firms within the same group. If $\sigma = 0$ then the logit model obtains.

For single-product firms, the first-order condition for profit maximization $p_{jt} - mc_{jt} = -\frac{1}{dq_{jt}/dp_{jt}} q_{jt}$ can be expressed as

$$\mu_{jt} = -\frac{1}{ds_{jt}/dp_{jt}} s_{jt} \tag{A.3}$$

For equation (3) to hold, it must be that $\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) = -\alpha\mu_{jt}$. We can solve for $\lambda^{(j,t)}$ in the nested logit model by taking the derivatives of s with respect to p to obtain:

$$\lambda^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) = \frac{1}{\frac{1}{1-\sigma} - s_{jt} - \frac{\sigma}{1-\sigma}\bar{s}_{j|g,t}}. \tag{A.4}$$

Thus, $\lambda^{(j,t)}$ can be expressed as a function of $\boldsymbol{q}_t$, and $\boldsymbol{w}_t$, and θ_2. In our third application, we use the nested logit model to estimate bounds on the structural parameters (Section 5.3).

A.2 Random Coefficients Logit Demand

In our application in Section 5.1, we develop the underlying indirect utility model of the random coefficients logit model, following Berry (1994) and Nevo (2000, 2001). Here, we provide some additional results using the notation from that section.

The probability with which consumer i selects product j ($j \neq 0$) is

$$\rho_{ijt}(\boldsymbol{\delta}_t, \boldsymbol{w}_t; \theta_2) \equiv \frac{\exp(\delta_{jt} + \phi_{ijt}(\boldsymbol{w}_t; \theta_2))}{1 + \sum_{k=1}^{J} \exp(\delta_{kt} + \phi_{ikt}(\boldsymbol{w}_t; \theta_2))} \tag{A.5}$$

which is obtained under the assumption that ϵ is distributed i.i.d. type 1 extreme value. Equation (30) aggregates choice probabilities across consumers. This calculation of expected market shares converges to observed shares as $N_t \to \infty$. In implementation, equation (30) is often approximated by summing across a number of simulated consumers, with each simulated consumer being characterized by a set of demographics $\{D_i, \nu_i\}$.

Consider a candidate parameter vector $\tilde{\theta}_2$ that includes Π and Σ. Given $\tilde{\theta}_2$, Berry et al. (1995) prove that a contraction mapping recovers the $J \times 1$ vector $\boldsymbol{\delta}_t(\boldsymbol{s}_t, \boldsymbol{w}_t, \tilde{\theta}_2)$ that such that the choice probabilities implied by the model match the market shares observed in the data. This "mean valuation" vector is equivalent to the vector $\boldsymbol{h}_t$ in our notation. These vectors can be stacked to obtain the full $JT \times 1$ vector $\boldsymbol{H}$ in a single procedure.

The supply restriction from equation (3) is satisfied when (multi-product) firms compete by setting prices, following our more general results for differentiated products Bertrand in Appendix A.5. For example, for the special case with single-product firms and no random coefficient on price ($\alpha = \alpha_i \forall i$), the Bertrand-Nash equilibrium yields

$$\lambda^{(j,t)}(\boldsymbol{s}_t, \boldsymbol{w}_t; \theta_2) = \frac{s_{jt}}{\frac{1}{N_t}\sum_i \rho_{ijt}(1 - \rho_{ijt})} \tag{A.6}$$

where the denominators integrate over the (product of) consumer-specific choice probabilities. From an econometric standpoint, λ is free from the price parameter α because it depends only on market shares and consumer-specific choice probabilities, $\rho_{ijt}(\boldsymbol{\delta}_t, \boldsymbol{w}_t; \theta_2)$. As discussed above, $\boldsymbol{\delta}_t$ can obtained as a function of shares.

More complicated versions of $\lambda^{(j,t)}$ can be constructed numerically, however, this step is not necessary as estimation can proceed by implementing the covariance restriction directly using the method of moments. Confirming the restrictions on h and λ ensures identification of the price coefficient and other linear parameters, conditional on $\tilde{\theta}_2$. The identification of θ_2 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 the application in Section 5.1.

A.3 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:

$$h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) \equiv \ln(q_{jt}/\overline{q}_t) = \beta + \alpha \ln\left(\frac{p_{jt}}{\Pi_t}\right) + \xi_{jt}$$

where $\overline{q}_t$ is an observed demand shifter, Π_t is a price index, and α provides the constant elasticity of demand. In our notation, $\overline{q}_t$ and Π_t are contained in $\boldsymbol{w}_t$.

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 when marginal costs are constant.[33] 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.

A.4 Other Demand Systems

The demand assumption in equation (1) accommodates many rich demand systems. Consider the linear demand system, $q_{jt} = \beta_j + \sum_k \alpha_{jk} p_k + \xi_{jt}$, which sometimes appears in identification proofs (e.g., Nevo, 1998) but is seldom applied empirically due to the large number of price coefficients.

The system can be formulated such that $h^{(j,t)}(\boldsymbol{q}_t, \boldsymbol{w}_t; \theta_2) \equiv q_{jt} - \sum_{k \neq j} \alpha_{jk} p_{kt}$. In this demand system, other prices (p_{kt}) are elements of $\boldsymbol{w}_t$ and the cross-product price coefficients $\alpha_{jk}(k \neq j)$ are elements of θ_2. In addition to the own-product uncorrelatedness restrictions that could identify α_{jj}, one could impose cross-product covariance restrictions to identify α_{jk} ($k \neq j$). We discuss these cross-product covariance restrictions in the first application (Section 5.1). A similar approach could be used with the almost ideal demand system of Deaton and Muellbauer (1980).

A.5 Multi-Product Firms with Bertrand Competition

We illustrate how our framework more generally incorporates multi-product firms with the case of Bertrand pricing. For this setting, we assume that the derivatives $\partial q_j / \partial p_k$ exist and that $\boldsymbol{D}$,

[33] 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 $J \times J$ matrix of derivatives $\partial q_j / \partial p_k \ \forall j, k \in \mathcal{J}$, is invertible. For matrix elements, let j index rows and k index columns. The market subscript, t, is omitted to simplify notation.

We begin by establishing properties of demand using the restriction $h^{(j)}(\boldsymbol{q}, \boldsymbol{w}; \theta_2) = \alpha p_j + x_j \beta + \xi_j$ from equation (1). Taking the derivative with respect to p_j (holding fixed $\boldsymbol{x}$, $\boldsymbol{\xi}$, and $p_{kt} \ \forall k \neq j$), we obtain

$$\sum_k \frac{\partial h^{(j)}}{\partial q_k} \frac{\partial q_k}{\partial p_j} = \alpha \tag{A.7}$$

Similarly, we obtain $\sum_k \frac{\partial h^{(j)}}{\partial q_k} \frac{\partial q_k}{\partial p_\ell} = 0 \ \forall \ell \neq j$. These restrictions on demand admit the expression

$$\boldsymbol{\mathcal{H}} \boldsymbol{D} = \alpha \boldsymbol{I} \implies \boldsymbol{D} = \alpha \boldsymbol{\mathcal{H}}^{-1} \tag{A.8}$$

where $\boldsymbol{\mathcal{H}}$ denotes the $J \times J$ matrix of derivatives $\partial h^{(j)} / \partial q_k \ \forall j, k \in \mathcal{J}$, $\boldsymbol{D}$ is defined as above, and $\boldsymbol{I}$ is a $J \times J$ identity matrix. Because $h^{(j)}$ is a known function of $\boldsymbol{q}$, $\boldsymbol{w}$, and θ_2, its derivatives with respect to q_k and thus each element of $\boldsymbol{\mathcal{H}}$ can be calculated, and $\boldsymbol{\mathcal{H}}^{-1}$ can be solved for. We let $\boldsymbol{A} \equiv \boldsymbol{\mathcal{H}}^{-1}$ such that $\boldsymbol{A}$ denotes the matrix of demand derivatives up to the scalar α.

We now turn to the supply side of the model. 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

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

Let $\boldsymbol{D}^m$ denote the matrix of derivatives $\partial q_j / \partial p_k \ \forall j, k \in K^m$, and let $\boldsymbol{q}^m$ and $\boldsymbol{\mu}^m$ denote the stacked vector of q_j and $\mu_j = p_j - mc_j$ for $j \in K^m$. Stacking the first-order conditions yields

$$\boldsymbol{D}^m \boldsymbol{\mu}^m = -\boldsymbol{q}^m \tag{A.9}$$

and, solving for markups, $\boldsymbol{\mu}^m$, we obtain

$$\boldsymbol{\mu}^m = -\left(\boldsymbol{D}^m\right)^{-1} \boldsymbol{q}^m \tag{A.10}$$

Equation (3) requires the construction of λ such that $\boldsymbol{\mu}^m = -\frac{1}{\alpha} \boldsymbol{\lambda}^m$. In the case of multi-product firms with Bertrand competition, it immediately follows that

$$\boldsymbol{\lambda}^m = \alpha \left(\boldsymbol{D}^m\right)^{-1} \boldsymbol{q}^m \tag{A.11}$$

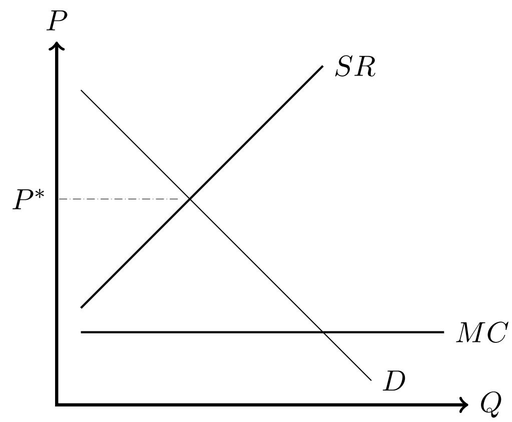
Following the conditions on demand above, we have $\boldsymbol{D} = \alpha \boldsymbol{A}$, and taking the corresponding element-by-element minor, we have $\boldsymbol{D}^m = \alpha \boldsymbol{A}^m$. This yields

$$\boldsymbol{\lambda}^m = \alpha \left(\alpha \boldsymbol{A}^m\right)^{-1} \boldsymbol{q}^m \tag{A.12}$$

$$\boldsymbol{\lambda}^m = \left(\boldsymbol{A}^m\right)^{-1} \boldsymbol{q}^m \tag{A.13}$$

Because $\boldsymbol{A}^m$ is a known function of $\boldsymbol{q}$, $\boldsymbol{w}$, and θ_2, $\boldsymbol{\lambda}^m$ can be constructed of the same arguments. Thus, we see that multi-product Bertrand fits in the class of models specified by equation (3).

Figure A.1: Supply Relationship



Notes: Figure plots an illustrative example of demand (D), marginal costs (MC), and the supply relationship described in the paper (SR). The supply relationship can be interpreted as the opportunity cost to the firm of selling an additional unit. The opportunity cost is the sum of the marginal cost and the inframarginal losses of lowering price. The equilibrium price (P^*) is determined by the intersection of D and SR.

A.6 Alternative Models of Competition

Our restriction on additive markups from equation (3) applies to a broad 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:

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

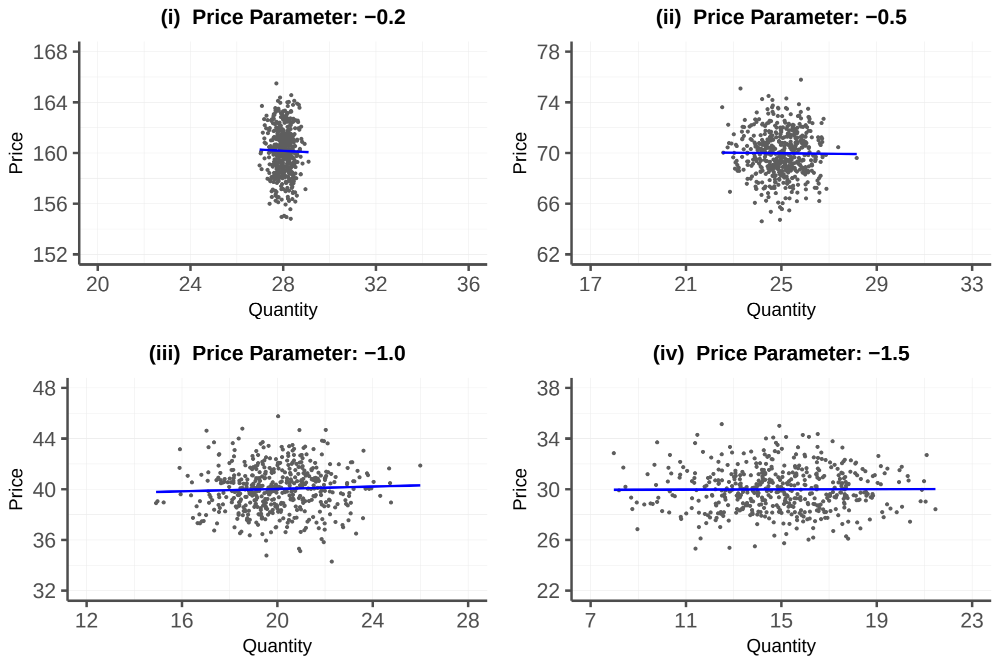
This generalized model of Nash competition nests Bertrand ($a = p$) and Cournot ($a = q$). The first-order condition, holding fixed the actions of the other firms, is given by:

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

In equilibrium, we obtain the structural decomposition $p = mc + \mu$, where μ incorporates the structure of demand and its parameters. This decomposition provides a restriction on how prices move with demand shocks, aiding identification. Using restrictions about demand, such as those imposed by equation (1), one can construct the appropriate form of $\lambda^{(j,t)}(\cdot)$ and solve for the price coefficient. Related first-order conditions can be obtained in other contexts, such as consistent conjectures.

Bresnahan (1982) refers to the above equation as the "supply relation," and notes that it generalizes to many different forms of conduct. Figure A.1 plots the supply relationship along with the demand curve for an illustrative setting. The supply relationship lies above the marginal cost curve, and the difference is given by the inframarginal loss in revenue for selling an additional unit (i.e., the gap between price and marginal revenue). As the inframarginal loss has an opportunity cost interpretation, the supply relation can be conceptualized as the sum of the marginal cost curve and the firm's opportunity cost curve, with the latter incorporating any market power that the firm has. Equilibrium price is determined by the intersection of demand

Figure B.1: Relative Variation in Prices and Quantities with Different Price Parameters



Notes: This figure displays equilibrium prices and quantities with four different values of the price parameter. The supports of $\Delta\xi$ and $\Delta\eta$ are selected so that the supply-side and demand-side variation is balanced and that the equilibrium price-quantity pairs form a cloud with no apparent slope. The line in each figure indicates the slope obtained by OLS regression.

and the supply relationship. This is equivalent to the equilibrium price that obtains if the firm sets price to equate marginal revenue and marginal cost.

B Empirical Variation in Prices and Quantities

In Section 2.3, we state that the empirical variation that identifies the price parameter (α) under a covariance restriction relates to the relative variation in (transformed) quantities and prices. In this appendix, we provide a numerical example for illustrative purposes using the monopoly model of equations (18) and (19).

We assume that demand and marginal cost are given by

$$q_{jt} = 60 + \alpha p_{jt} + \Delta\xi_{jt} \qquad \text{and} \qquad mc_{jt} = 20 + \Delta\eta_{jt}$$

We consider four values of α: -0.2, -0.5, -1.0, and -1.5. As the slope of the inverse demand relationship is $1/\alpha$ and that of the inverse supply relationship is $-1/\alpha$, the more negative values for α generate flatter inverse demand/supply relationships. We let $\Delta\eta \sim 2.5 \times N(0,1)$ and let $\Delta\xi \sim \frac{2.5}{-\alpha} \times N(0,1)$ with α affecting the support of $\Delta\xi$ so that both variables have the same support if measured in the same units. We then take 500 draws on these demand and cost shocks for each of the four price parameters, and compute the equilibrium prices and quantities.

Figure B.1 shows the results. The four panels correspond to the four values of α. The panels have comparable scales (16 units by 16 units) and are re-centered along the x and y axes. In each, the equilibrium price-quantity pairs form a cloud with no apparent slope. The reason is that supports of $\Delta\xi$ and $\Delta\eta$ are selected so that the supply-side and demand-side variation is balanced. (Figure 1 illustrates how the clouds would slope down if the supply shocks dominate, and up if the demand shocks dominate.) The four panels in Figure B.1 illustrate that a more negative value of α leads to greater variation in quantities relative to variation in prices. The reason is that the inverse demand and supply relationships are flatter, and uncorrelated shifts in flatter inverse supply and demand relationships produce more variation in quantities, all else equal. Intuitively then, it should be possible to compare the relative variance of quantity and price to learn about the price coefficient. Proposition 3 formalizes this result.

For the monopoly model that we use for this numerical example—which features linear demand and constant marginal cost—the approximation provided in Proposition 3 is exact and simplifies to $\alpha^{CR} = -\sqrt{Var(q)}/\sqrt{Var(p)}$. Calculating the implied estimate for each scatter plot, we obtain -0.19, -0.50, -0.98, and -1.54, which are close to respective values of the parameters used to generate the data.

C Supply-Side Misspecification

To illustrate how supply-side misspecification can 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.[34] We assume the demand system is logit, providing consumers with a differentiated discrete choice, and we allow them to select an outside option in addition to a product from each firm. The quantity demanded of firm j in market t is

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

On the supply side, marginal costs are $c_{kt} = \Delta\eta_{kt}$ $(k = j, i)$. 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.

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. We then estimate the model using the covariance restrictions approach assuming $Cov(\Delta\xi, \Delta\eta) = 0$, using $\Delta\eta$ as an (observed) excluded instrument for demand, and using $\Delta\xi$ as an (observed) excluded instrument for supply. Across all

[34] Another form of misspecification could arise if prices or quantities are measured with error, in which case the demand and cost residuals might be correlated even if the underlying shocks are uncorrelated.

Table C.1: 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 $h_{jt} = 2 - p_{jt} + \Delta\xi_{jt}$, so that $\alpha = -1$, and marginal costs are $c_{jt} = \Delta\eta_{jt}$. 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 ($\Delta\eta$) as an instrument in the demand equation. Analogously, IV-2 estimates are calculated using two-stage least squares with demand shocks ($\Delta\xi$) as an instrument in the supply relationship. Across all specifications, $\Delta\xi \sim U(0, 0.5)$ and $\Delta\eta \sim U(0, 0.5)$. 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 ($\kappa = 0$).

specifications, $\Delta\xi \sim U(0, 0.5)$ and $\Delta\eta \sim U(0, 0.5)$.[35]

Table C.1 displays the results. As expected, supply-side misspecification can introduce bias into the covariance restrictions approach. The bias does not appear to be meaningful for modest values of κ (i.e., 0.6 or less). When the true nature of conduct is $\kappa = 1$ (joint price setting), but we assume Bertrand price competition, the bias is -3.8 percent. 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 appears to partially mitigate the bias from model misspecification.

[35] We note that these are not mean zero, but it does not matter in this case. It is simply a normalization.

D Proofs

D.1 A Consistent and Unbiased Quasi-Estimator of $\Delta\xi$

Our proofs make use of the following lemma, which identifies a consistent and unbiased quasi-estimator for the unobserved term in a linear regression when one of the covariates is endogenous. We refer to it as a *quasi-estimator* because it depends on unobservables and cannot be constructed from the data. It turns out to be a useful input in our proofs. 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.

For convenience, in this section, we omit the market-period subscripts jt on scalar variables such as p, h, and ξ and the $K \times 1$ (row) vector x.

Lemma D.1. *A consistent and unbiased quasi-estimator of $\Delta\xi$ is given by $\hat{\Delta\xi}_1 = \hat{\Delta\xi}^{OLS} + \left(\hat{\alpha}^{OLS} - \alpha\right) p^*$*

For some intuition, note that we can construct both the true demand shock and OLS residuals (at the probability limit) as:

$$\Delta\xi = h - \alpha p - x\beta$$
$$\Delta\xi^{OLS} = h - \alpha^{OLS} p - x\beta^{OLS}$$

where this holds even in small samples. Recall that $E[\Delta\xi_{jt}] = 0$. The true demand shock is given by $\Delta\xi_0 = \Delta\xi^{OLS} + (\alpha^{OLS} - \alpha)p + x(\beta^{OLS} - \beta)$.

We desire to show that the quasi-estimator of the demand shock, $\hat{\Delta\xi}_1 = \hat{\Delta}\xi^{OLS} + \left(\hat{\alpha}^{OLS} - \alpha\right) p^*$, is consistent and unbiased. This eliminates the need to estimate the true β parameters. It suffices to show that $(\hat{\alpha}^{OLS} - \alpha)p^* = (\hat{\alpha}^{OLS} - \alpha)p + x(\hat{\beta}^{OLS} - \beta) + \Upsilon$, where Υ is such that $E[\Upsilon = 0]$ and $\Upsilon \to 0$ as T gets large. It is straightforward to show this using the projection matrices for p and x.[36]

D.2 Proof of Proposition 1 (Set Identification)

From equation (9), we have $\hat{\alpha}^{OLS} \xrightarrow{p} \alpha + \frac{Cov(p^*, \Delta\xi)}{Var(p^*)}$. The general form for a firm's first-order condition is $p = mc + \mu$, where mc is the marginal cost and μ is the markup. We can write $p = p^* + \hat{p}$, where $\hat{p}$ is the projection of p on the exogenous variables $\tilde{x}$ that include product and market fixed effects. If we substitute the first-order condition $p^* = mc + \mu - \hat{p}$ into the bias term from the OLS regression, we obtain

$$\begin{aligned} \alpha^{OLS} - \alpha = \frac{Cov(p^*, \Delta\xi)}{Var(p^*)} &= \frac{Cov(\Delta\xi, mc + \mu - \hat{p})}{Var(p^*)} \\ &= \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} + \frac{Cov(\Delta\xi, \mu)}{Var(p^*)} \end{aligned} \tag{D.1}$$

[36] Please contact the authors if interested in the full proof.

where the second line follows from the exogeneity assumption that $E[\Delta\xi|\tilde{\boldsymbol{x}}] = 0$ and that, by assumption, $mc = x\gamma + \eta$. The exogeneity assumption implies that $\Delta\xi$ is orthogonal to the product-specific and time-specific terms in mc, as these are included in $\tilde{x}$ as fixed effects.

From Lemma D.1, we can construct a consistent estimate of the unobserved demand shock as $\Delta\xi = \Delta\xi^{OLS} + \left(\alpha^{OLS} - \alpha\right) p^*$. We substitute this expression into $\frac{Cov(\Delta\xi,\mu)}{Var(p^*)}$, along with the above expression for $\left(\alpha^{OLS} - \alpha\right)$ to obtain

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

Plugging this into equation (D.1) yields

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

Thus, we obtain an expression for the plim of 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 as a function of observable data, and if the correlation in unobserved shocks can be calibrated, we have a method to estimate α from the OLS regression. Under our supply and demand assumptions, $\mu = -\frac{1}{\alpha}\lambda$, and plugging in obtains the first equation of the proposition:

$$\alpha^{OLS} = \alpha - \frac{1}{\alpha + \frac{Cov(p^*,\lambda)}{Var(p^*)}}\frac{Cov\left(\Delta\xi^{OLS},\lambda\right)}{Var(p^*)} + \alpha\frac{1}{\alpha + \frac{Cov(p^*,\lambda)}{Var(p^*)}}\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}.$$

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

D.3 Proof of Proposition 2 (Point Identification)

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

$$0 \leq \alpha^{OLS}\frac{Cov(p^*,\lambda)}{Var(p^*)} + \frac{Cov\left(\Delta\xi^{OLS},\lambda\right)}{Var(p^*)} \tag{D.2}$$

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

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

Lemma D.2. *The roots of equation* (11) *are* α *and* $\frac{Cov(p^*,\Delta\xi)}{Var(p^*)} - \frac{Cov(p^*,\lambda)}{Var(p^*)} - \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}$.

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

$$\begin{aligned} 0 &= \alpha^2 \\ &+ \left(\frac{Cov(p^*,\lambda)}{Var(p^*)} + \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} - \alpha^{OLS}\right)\alpha \\ &+ \left(-\alpha^{OLS}\frac{Cov(p^*,\lambda)}{Var(p^*)} - \frac{Cov\left(\Delta\xi^{OLS},\lambda\right)}{Var(p^*)}\right) \end{aligned}$$

To find the roots, begin by applying the quadratic formula

$$\begin{aligned} (r_1, r_2) &= \frac{1}{2}\left(-B \pm \sqrt{B^2 - 4AC}\right) \\ &= \frac{1}{2}\left(\alpha^{OLS} - \frac{Cov(p^*,\lambda)}{Var(p^*)} - \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}\right) \\ &\pm\frac{1}{2}\sqrt{\left(\alpha^{OLS} + \frac{Cov(p^*,\lambda)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\Delta\xi^{OLS},\lambda\right)}{Var(p^*)} + \left(\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}\left(\alpha^{OLS} - \frac{Cov(p^*,\lambda)}{Var(p^*)}\right)} \end{aligned} \tag{D.3}$$

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

$$\begin{aligned} &\left(\alpha^{OLS} + \frac{Cov(p^*,\lambda)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\Delta\xi^{OLS},\lambda\right)}{Var(p^*)} \\ &= \left(\alpha^{OLS} + \frac{Cov(p^*,\lambda)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\Delta\xi - p^*(\alpha^{OLS} - \alpha),\lambda\right)}{Var(p^*)} \\ &= \left(\alpha^{OLS} + \frac{Cov(p^*,\lambda)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\Delta\xi,\lambda\right)}{Var(p^*)} - 4\frac{Cov(p^*,\Delta\xi)}{Var(p^*)}\frac{Cov\left(p^*,\lambda\right)}{Var(p^*)} \\ &= \left(\alpha^{OLS} + \frac{Cov(p^*,\lambda)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\Delta\xi,\lambda\right)}{Var(p^*)} - 4\left(\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} + \frac{Cov(\Delta\xi,-\frac{1}{\alpha}\lambda)}{Var(p^*)}\right)\frac{Cov\left(p^*,\lambda\right)}{Var(p^*)} \\ &= \left(\alpha^{OLS} + \frac{Cov(p^*,\lambda)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\Delta\xi,\lambda\right)}{Var(p^*)}\left(1 + \frac{1}{\alpha}\frac{Cov\left(p^*,\lambda\right)}{Var(p^*)}\right) - 4\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}\frac{Cov\left(p^*,\lambda\right)}{Var(p^*)} \end{aligned} \tag{D.4}$$

To simplify this expression, it is helpful to use the general decomposition of a firm's first-order condition, $p = mc + \mu$, where mc 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, $\tilde{x}$. By assumption, $c = x\gamma + \eta$. It follows that

$$\begin{aligned} p^* &= x\gamma + \eta + \mu - \hat{p} \\ &= x\gamma + \eta - \frac{1}{\alpha}\lambda - \hat{p} \end{aligned}$$

Therefore

$$Cov(p^*, \Delta\xi) = Cov(\Delta\xi, \Delta\eta) - \frac{1}{\alpha}Cov(\Delta\xi, \lambda)$$

and

$$\begin{aligned}
Cov(\Delta\xi, \lambda) &= -\alpha \left(Cov(p^*, \Delta\xi) - Cov(\Delta\xi, \Delta\eta)\right) \\
\frac{Cov(\Delta\xi, \lambda)}{Var(p^*)} &= -\alpha \left(\frac{Cov(p^*, \Delta\xi)}{Var(p^*)} - \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right)
\end{aligned} \tag{D.5}$$

Returning to equation (D.4), we can substitute using equation (D.5) and simplify:

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

Now, consider the second part inside of the radical in equation (D.3):

$$\begin{aligned}
&\left(\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\left(\alpha^{OLS} - \frac{Cov(p^*, \lambda)}{Var(p^*)}\right) \\
=& \left(\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\left(\alpha + \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} - \frac{1}{\alpha}\frac{Cov(\Delta\xi, \lambda)}{Var(p^*)} - \frac{Cov(p^*, \lambda)}{Var(p^*)}\right) \\
=& \left(\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right)^2 - 2\alpha\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} - 2\left(\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right)^2 + 2\frac{1}{\alpha}\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\frac{Cov(\Delta\xi, \lambda)}{Var(p^*)} + 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\frac{Cov(p^*, \lambda)}{Var(p^*)} \\
=& - \left(\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right)^2 - 2\alpha\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} - 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\left(\frac{Cov(p^*, \Delta\xi)}{Var(p^*)} - \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right) + 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\frac{Cov(p^*, \lambda)}{Var(p^*)} \\
=& \left(\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\right)^2 - 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\alpha - 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\frac{Cov(p^*, \Delta\xi)}{Var(p^*)} + 2\frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)}\frac{Cov(p^*, \lambda)}{Var(p^*)}
\end{aligned}$$

Combining yields a simpler expression for the terms inside the radical of equation (D.3):

$$\begin{aligned}
&\left(\left(\alpha+\frac{Cov\left(p^{*},\lambda\right)}{Var(p^{*})}\right)-\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}\right)^{2}+4\alpha\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\\
&+\left(\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\right)^{2}-2\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\alpha-2\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\frac{Cov(p^{*},\Delta\xi)}{Var(p^{*})}+2\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\frac{Cov(p^{*},\lambda)}{Var(p^{*})}\\
=&\left(\left(\alpha+\frac{Cov\left(p^{*},\lambda\right)}{Var(p^{*})}\right)-\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}\right)^{2}+\left(\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\right)^{2}\\
&+2\alpha\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}-2\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\frac{Cov(p^{*},\Delta\xi)}{Var(p^{*})}+2\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\frac{Cov(p^{*},\lambda)}{Var(p^{*})}\\
=&\left(\alpha+\frac{Cov\left(p^{*},\lambda\right)}{Var(p^{*})}-\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}+\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\right)^{2}
\end{aligned}$$

Plugging this back into equation (D.3), we have:

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

The roots are given by

$$\begin{aligned}
&\frac{1}{2}\left(\alpha+\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}-\frac{Cov(p^{*},\lambda)}{Var(p^{*})}-\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\right)+\\
&\qquad\frac{1}{2}\left(\alpha+\frac{Cov\left(p^{*},\lambda\right)}{Var(p^{*})}-\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}+\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\right)=\alpha
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{2}\left(\alpha+\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}-\frac{Cov(p^{*},\lambda)}{Var(p^{*})}-\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\right)+\\
&\qquad\frac{1}{2}\left(-\alpha-\frac{Cov\left(p^{*},\lambda\right)}{Var(p^{*})}+\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}-\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}\right)\\
&\qquad\qquad=\frac{Cov\left(p^{*},\Delta\xi\right)}{Var(p^{*})}-\frac{Cov(p^{*},\lambda)}{Var(p^{*})}-\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}
\end{aligned}$$

which completes the proof of the intermediate result. QED.

Part (3). Consider the roots of equation (11), α and $\frac{Cov(p^{*},\Delta\xi)}{Var(p^{*})}-\frac{Cov(p^{*},\lambda)}{Var(p^{*})}-\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^{*})}$. The price parameter α may or may not be the lower root.[37] However, α is the lower root iff

$$\begin{aligned}
\alpha &< \frac{Cov(p^*,\Delta\xi)}{Var(p^*)} - \frac{Cov(p^*,\lambda)}{Var(p^*)} - \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} \\
\alpha &< -\alpha\frac{Cov(p^*,-\frac{1}{\alpha}\Delta\xi)}{Var(p^*)} + \alpha\frac{Cov(p^*,-\frac{1}{\alpha}\lambda)}{Var(p^*)} - \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} \\
\alpha &< -\alpha\frac{Cov(p^*,-\frac{1}{\alpha}\Delta\xi)}{Var(p^*)} + \alpha\frac{Cov(p^*,p^*-c)}{Var(p^*)} - \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} \\
\alpha &< \alpha\frac{Var(p^*)}{Var(p^*)} - \alpha\frac{Cov(p^*,-\frac{1}{\alpha}\Delta\xi)}{Var(p^*)} - \alpha\frac{Cov(p^*,\Delta\eta)}{Var(p^*)} - \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} \\
0 &< -\alpha\frac{Cov(p^*,-\frac{1}{\alpha}\Delta\xi)}{Var(p^*)} - \alpha\frac{Cov(p^*,\Delta\eta)}{Var(p^*)} - \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} \\
0 &< \frac{Cov(p^*,-\frac{1}{\alpha}\Delta\xi)}{Var(p^*)} + \frac{Cov(p^*,\Delta\eta)}{Var(p^*)} + \frac{1}{\alpha}\frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}
\end{aligned}$$

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

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

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

D.4 Proof of Proposition 3 (Approximation)

The demand and supply equations are given by:

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

where $\frac{dh}{dq}q = \lambda$ for single-product firms. For ease of exposition, here we slightly abuse notation and assume that ξ and η are exogenous (and x includes dummy variables to absorb fixed

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

$$\alpha + \frac{Cov\left(p^*,\lambda\right)}{Var(p^*)} - \frac{Cov\left(p^*,\Delta\xi\right)}{Var(p^*)} + \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)} > 0$$

because, in that case,

$$\sqrt{\left(\alpha + \frac{Cov\left(p^*,\lambda\right)}{Var(p^*)} - \frac{Cov\left(p^*,\Delta\xi\right)}{Var(p^*)} + \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}\right)^2} = \alpha + \frac{Cov\left(p^*,\lambda\right)}{Var(p^*)} - \frac{Cov\left(p^*,\Delta\xi\right)}{Var(p^*)} + \frac{Cov(\Delta\xi,\Delta\eta)}{Var(p^*)}$$

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

effects). 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

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

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

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

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

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

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

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

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

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

$$h^* \approx \frac{1}{2}\left(\alpha\eta + \xi - H^*\right). \tag{D.7}$$

Equations (D.6) and (D.7) provide an approximation for α.

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

QED.

D.5 Proof of Lemma 1 (Monotonicity in $Cov(\Delta\xi, \Delta\eta)$)

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 $\gamma = Cov(\Delta\xi, \Delta\eta)$. 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

$$1 + \frac{b}{(b^2 - 4c)^{\frac{1}{2}}} > 0. \tag{D.8}$$

We have two cases. First, when $c < 0$, we know that $\left|\frac{b}{(b^2-4c)^{\frac{1}{2}}}\right| < 1$, which satisfies equation (D.8). 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 equation (D.8) 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

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

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 $[-\frac{1}{2}b, 0)$.

D.6 Proof of Proposition 4 (Covariance Bound)

The proof involves 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 equation (11), real solutions satisfy the condition:

$$\left(\frac{Cov(p^*, \lambda)}{Var(p^*)} + \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} - \alpha^{OLS}\right)^2 \geq 4\left(-\alpha^{OLS}\frac{Cov(p^*, \lambda)}{Var(p^*)} - \frac{Cov\left(\Delta\xi^{OLS}, \lambda\right)}{Var(p^*)}\right).$$

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

$$\left(\frac{Cov(p^*, \lambda)}{Var(p^*)} + \frac{Cov(\Delta\xi, \Delta\eta)}{Var(p^*)} - \alpha^{OLS}\right) \geq 2\sqrt{-\alpha^{OLS}\frac{Cov(p^*, \lambda)}{Var(p^*)} - \frac{Cov\left(\Delta\xi^{OLS}, \lambda\right)}{Var(p^*)}}.$$

Solving for $Cov(\Delta\xi, \Delta\eta)$, we obtain the model-based bound,

$$Cov(\Delta\xi, \Delta\eta) \geq Var(p^*)\alpha^{OLS} - Cov(p^*, \lambda) + 2Var(p^*)\sqrt{-\alpha^{OLS}\frac{Cov(p^*, \lambda)}{Var(p^*)} - \frac{Cov\left(\Delta\xi^{OLS}, \lambda\right)}{Var(p^*)}}.$$

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.

D.7 Proof of Corollary 1 (Marginal Cost Functions)

Under the semi-linear marginal cost schedule of equation (26) and the assumption that $Cov(\Delta\xi, \Delta\eta) = 0$, the plim of the OLS estimator is equal to

$$\text{plim } \hat{\alpha}^{OLS} = \alpha + \frac{Cov(\Delta\xi, g)}{Var(p^*)} - \frac{1}{\alpha}\frac{Cov\left(\Delta\xi, \lambda\right)}{Var(p^*)}.$$

This is obtain directly by plugging in the first–order condition for p: $Cov(p^*, \Delta\xi) = Cov(g(q;\tau) + \eta - \frac{1}{\alpha}\lambda - \hat{p}, \Delta\xi) = Cov(\Delta\xi, g) - \frac{1}{\alpha}Cov(\Delta\xi, \lambda)$ under the assumptions. Next, we re-express the terms including the unobserved demand shocks in in terms of OLS residuals. As shown by Lemma D.1, the estimated residuals are given by $\Delta\xi^{OLS} = \Delta\xi + \left(\alpha - \alpha^{OLS}\right)p^*$. As $\alpha - \alpha^{OLS} = \frac{1}{\alpha}\frac{Cov(\Delta\xi,\lambda)}{Var(p^*)} - \frac{Cov(\Delta\xi,g)}{Var(p^*)}$, we obtain $\Delta\xi^{OLS} = \Delta\xi + \left(\frac{1}{\alpha}\frac{Cov(\Delta\xi,\lambda)}{Var(p^*)} - \frac{Cov(\Delta\xi,g)}{Var(p^*)}\right)p^*$. This implies

$$Cov\left(\Delta\xi^{OLS}, \lambda\right) = \left(1 + \frac{1}{\alpha}\frac{Cov(p^*, \lambda)}{Var(p^*)}\right)Cov(\Delta\xi, \lambda) - \frac{Cov(p^*, \lambda)}{Var(p^*)}Cov(\Delta\xi, g)$$
$$Cov\left(\Delta\xi^{OLS}, g(q;\tau)\right) = \frac{1}{\alpha}\frac{Cov(p^*, g)}{Var(p^*)}Cov\left(\Delta\xi, \lambda\right) + \left(1 - \frac{Cov(p^*, g)}{Var(p^*)}\right)Cov(\Delta\xi, g)$$

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

$$\begin{bmatrix} Cov(\Delta\xi,\lambda) \\ Cov(\Delta\xi,g) \end{bmatrix} = \begin{bmatrix} 1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)} & -\frac{Cov(p^*,\lambda)}{Var(p^*)} \\ \frac{1}{\alpha}\frac{Cov(p^*,g)}{Var(p^*)} & 1-\frac{Cov(p^*,g)}{Var(p^*)} \end{bmatrix}^{-1} \begin{bmatrix} Cov(\Delta\xi^{OLS},\lambda) \\ Cov(\Delta\xi^{OLS},g) \end{bmatrix}$$

where

$$\begin{bmatrix} 1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)} & -\frac{Cov(p^*,\lambda)}{Var(p^*)} \\ \frac{1}{\alpha}\frac{Cov(p^*,g)}{Var(p^*)} & 1-\frac{Cov(p^*,g)}{Var(p^*)} \end{bmatrix}^{-1} = \\ \frac{1}{1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)}-\frac{Cov(p^*,g)}{Var(p^*)}} \begin{bmatrix} 1-\frac{Cov(p^*,g)}{Var(p^*)} & \frac{Cov(p^*,\lambda)}{Var(p^*)} \\ -\frac{1}{\alpha}\frac{Cov(p^*,g)}{Var(p^*)} & 1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)} \end{bmatrix}.$$

Therefore, we obtain the relations

$$\begin{aligned} Cov(\Delta\xi,\lambda) &= \frac{\left(1-\frac{Cov(p^*,g)}{Var(p^*)}\right)Cov(\Delta\xi^{OLS},\lambda)+\frac{Cov(p^*,\lambda)}{Var(p^*)}Cov(\Delta\xi^{OLS},g)}{1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)}-\frac{Cov(p^*,g)}{Var(p^*)}} \\ Cov(\Delta\xi,g) &= \frac{-\frac{1}{\alpha}\frac{Cov(p^*,g)}{Var(p^*)}Cov(\Delta\xi^{OLS},\lambda)+\left(1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)}\right)Cov(\Delta\xi^{OLS},g)}{1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)}-\frac{Cov(p^*,g)}{Var(p^*)}}. \end{aligned}$$

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

$$\begin{aligned} &\left(1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)}-\frac{Cov(p^*,g)}{Var(p^*)}\right)\left(Cov(\Delta\xi,g)-\frac{1}{\alpha}Cov\left(\Delta\xi,\lambda\right)\right) \\ =&-\frac{1}{\alpha}\frac{Cov(p^*,g)}{Var(p^*)}Cov(\Delta\xi^{OLS},\lambda)+\left(1+\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)}\right)Cov(\Delta\xi^{OLS},g) \\ &-\frac{1}{\alpha}\left(1-\frac{Cov(p^*,g)}{Var(p^*)}\right)Cov(\Delta\xi^{OLS},\lambda)-\frac{1}{\alpha}\frac{Cov(p^*,\lambda)}{Var(p^*)}Cov(\Delta\xi^{OLS},g) \\ =&Cov(\Delta\xi^{OLS},g)-\frac{1}{\alpha}Cov(\Delta\xi^{OLS},\lambda). \end{aligned}$$

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

$$\text{plim}\hat{\alpha}^{OLS}=\alpha-\frac{\frac{Cov(\Delta\xi^{OLS},\lambda)}{Var(p^*)}-\alpha\frac{Cov(\Delta\xi^{OLS},g)}{Var(p^*)}}{\alpha+\frac{Cov(p^*,\lambda)}{Var(p^*)}-\alpha\frac{Cov(p^*,g)}{Var(p^*)}},$$

and the following quadratic α.

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

QED.