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

Alexander MacKay **Nathan H. Miller**
Harvard University[†] Georgetown University[‡]

April 15, 2022

Abstract

We consider the identification of empirical models of supply and demand with imperfect competition. As is well known, a supply-side instrument can resolve price endogeneity in demand estimation. We demonstrate that, under common assumptions, two other approaches also yield consistent estimates of the joint model: (i) a demand-side instrument, or (ii) a covariance restriction between unobserved demand and cost shocks. Covariance restrictions can obtain identification even the absence of instruments. Further, supply and demand assumptions alone may bound the structural parameters without additional restrictions. 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

1 Introduction

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

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

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

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

---

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

because the endogenous data are interpreted directly through the lens of the model.[2] At a high level, this approach relates to Petterson et al. (2021), who show how to bound structural parameters based on beliefs about the magnitudes of unobserved shocks.

The results address one of the largest obstacles to research in empirical economics—that of finding valid instruments—and may allow researchers to push ahead along new frontiers. For example, Döpper et al. (2021) use our covariance restriction approach to estimate demand and markups across more than 100 consumer product categories. In that study, product and time fixed effects absorb the correlations that pose the most obvious threats to validity, such as the possibility that higher quality products are more expensive to produce. The residual variation in marginal costs (the "cost shock") then can be conceptualized as incorporating the contribution of instruments used elsewhere in the literature, such as input prices fluctuations that affect products differentially (Backus et al., 2021). In a setting where finding valid instruments would be difficult, the covariance restriction approach provides a theory-based path to recovering causal parameters in the presence of price endogeneity.

With nested logit and random coefficient logit demand—two models that often are used in industrial organization—additional identifying restrictions are necessary to identify parameters that characterize the heterogeneity of consumer preferences. Some recent applications use so-called "micro-moments" constructed from the observed behavior of individual consumers (e.g., Backus et al., 2021; Döpper et al., 2021) or "second-choice" data on what consumers view as their next-best option (e.g., Grieco et al., 2021). These strategies identify the consumer heterogeneity parameters but do not resolve price endogeneity (Berry and Haile, 2020), so the covariance restriction we examine is a useful complement. Alternatively, if instruments constructed from the characteristics of competing products (e.g., Berry et al., 1995; Gandhi and Houde, 2020), then the covariance restriction can be incorporated using the generalized method of moments (GMM) as an additional identifying restriction.

The strategy of using supply-side restrictions to reduce identification requirements has parallels in a handful of other articles. A simple linear example is provided in Koopmans (1949). Leamer (1981) examines a linear model of perfect competition, and provides conditions under which the price parameters can be bounded using only the endogenous variation in prices and quantities. Feenstra (1994) considers the case of monopolistic competition with constant markups, and a number of application in the trade literature extend this approach (e.g., Broda and Weinstein, 2006, 2010; Soderbery, 2015).[3] Zoutman et al. (2018) return to perfect com-

---

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

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

petition and show that, under a standard assumption in models of taxation, both supply and demand can be estimated with exogenous variation in a single tax rate. Our research builds on these articles by focusing on imperfect competition with adjustable markups.

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

In addition to our formal analysis, we demonstrate how to apply the approach in three distinct empirical settings. In each, we discuss the credibility of covariance restrictions. The applications also illustrate useful extensions, such as how to generate additional covariance restrictions and how to account for increasing marginal costs.

We organize the paper as follows. Section 2 describes the three approaches to identification, using demand and supply assumptions that are commonly employed in empirical studies of imperfect competition. Section 3 develops an identification strategy using covariance restrictions alone. Section 4 provides numerical simulations to compare the three approaches to identification. In small samples, the covariance restriction approach performs well, even when an instrument-based approach suffers from the weak instruments problem. In Section 5, we present three empirical applications.

2 Model

2.1 Data-Generating Process

The model examines supply and demand in a number of markets that can be conceptualized as geographically or temporally distinct. In each market t, there is a set $\mathcal{J}_t = \{0, 1, \ldots, J_t\}$ products available for purchase. The market t is defined by $(\mathcal{J}_t, \boldsymbol{\chi}_t)$, where

$$\boldsymbol{\chi}_t = (\boldsymbol{p}_t, \boldsymbol{x}_t, \boldsymbol{\xi}_t, \boldsymbol{\eta}_t)$$

contains characteristics of the product and market. Among these, $\boldsymbol{p}_t = (p_{1t}, \ldots, p_{J_t t})$ are prices, $\boldsymbol{x}_t = (\boldsymbol{x}_{1t}, \ldots, \boldsymbol{x}_{J_t t})$, are non-price product and market characteristics that are observable to the econometrician, and $\boldsymbol{\xi}_t = (\xi_{1t}, \ldots, \xi_{J_t t})$ and $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{J_t t})$ are demand-side and supply-side structural error terms, respectively, that represent unobservable product-level or market-level characteristics. Let each $p_{jt}, \xi_{jt}, \eta_{jt} \in \mathbb{R}$, each $\boldsymbol{x}_{jt} \in \mathbb{R}^K$, and the support of χ_t be χ. We assume that non-price characteristics are exogenous, in that $\mathbb{E}[\xi_{jt}|\boldsymbol{x}_t] = \mathbb{E}[\eta_{jt}|\boldsymbol{x}_t] = 0$ for all $j = 1, \ldots, J_t$, but that prices may be correlated with the structural error terms. Without loss of

generality, we assume that $\mathcal{J}_t = \mathcal{J} = \{0, 1, \ldots, J\}$ going forward.

We place restrictions on both demand and supply. For demand, we assume that the quantity sold of each product is determined by $q_{jt} = \sigma_{jt}(\chi_t; \boldsymbol{\theta})$, where each $\sigma_{jt}(\cdot)$ is a differentiable, invertible demand function and $\boldsymbol{\theta}$ is a vector of parameters. We place the following restriction on inverse demand:

$$h_{jt}(\boldsymbol{q}_t, \boldsymbol{p}_t, \boldsymbol{x}_t, \boldsymbol{\xi}_t; \boldsymbol{\theta}) \equiv \sigma_{jt}^{-1}(\boldsymbol{q}_t, \boldsymbol{p}_t, \boldsymbol{x}_t, \boldsymbol{\xi}_t; \boldsymbol{\theta}) = \beta p_{jt} + \boldsymbol{x}_{jt}'\boldsymbol{\alpha} + \xi_{jt} \tag{1}$$

where $\boldsymbol{q}_t = (q_{1t}, \ldots, q_{Jt})$. We assume downward-sloping demand (i.e., $\beta < 0$). Models that satisfy this restriction are used regularly in the empirical literature of industrial organization. For example, with logit demand, we have $\sigma_{jt}^{-1}(\cdot) \equiv \ln(s_{jt}) - \ln(s_{0t})$, where we follow convention and use markets shares, s_{jt}, in place of quantities.

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

$$p_{jt} = \mu_{jt}(\chi_t; \boldsymbol{\theta}) + mc_{jt}(\chi_t; \boldsymbol{\theta}).$$

In our baseline model, we assume that marginal costs are a linear function of covariates and an additively separable cost shock:

$$mc_{jt}(\chi_t; \boldsymbol{\theta}) = \boldsymbol{x}_{jt}'\boldsymbol{\gamma} + \eta_{jt}. \tag{2}$$

Marginal costs can vary with quantities if demand and cost shocks are correlated (e.g., as in Berry et al., 1995; Ciliberto et al., 2021). If, instead, $Cov(\xi_{jt}, \eta_{jt}) = 0$ then marginal costs are constant. We later consider an augmented model in which non-constant marginal costs are incorporated more explicitly (Section 3.4).

Finally, consistent with a broad class of oligopoly models, we assume that additive markups take the form

$$\mu_{jt}(\chi_t; \boldsymbol{\theta}) = -\frac{1}{\beta}\lambda_{jt}(\boldsymbol{q}_t, \boldsymbol{p}_t, D(\chi_t), \boldsymbol{\eta}_t; \boldsymbol{\theta}), \tag{3}$$

where $D(\chi_t)$ denotes the $J \times J$ matrix of partial derivatives $\left[\frac{\partial \sigma_{kt}(\chi_t)}{\partial p_{lt}}\right]_{k,l}$. We show how to construct λ_{jt} from data in our applications and in Appendix A. For example: with single-product Bertrand pricing, we have $\mu_{jt}(\cdot) = -\frac{1}{dq_{jt}/dp_{jt}}q_{jt} = -\frac{1}{\beta}\frac{dh_{jt}}{dq_{jt}}q_{jt}$. If, in addition, demand is logit, then $\lambda_{jt}(\cdot) \equiv \frac{dh_{jt}}{dq_{jt}}q_{jt} = \frac{1}{1-s_{jt}}$.

The restrictions above yield the following inverse supply equation:

$$\lambda_{jt}(\boldsymbol{q}_t, \boldsymbol{p}_t, D(\chi_t); \boldsymbol{\theta}) = -\beta p_{jt} + \beta \boldsymbol{x}_{jt}'\boldsymbol{\gamma} + \beta \eta_{jt} \tag{4}$$

Supply is upward sloping (as $\beta < 0$) due to market power: higher prices are needed to induce firms to supply greater quantities. These markup adjustments imply that prices respond to demand shocks, even if marginal costs are constant in output.

Together, equations (1) and (4) constitute a system of supply and demand equations that characterizes equilibrium for widely-used empirical oligopoly models. The equations obtain, for example, with multi-product Bertrand pricing and either nested logit, random coefficients logit, or constant elasticity demands, as well as with some models of Cournot competition and collusion. Thus, the results we obtain are general to many empirical settings. In Appendix A, we provide typical examples with different assumptions about competition and demand.

2.2 Three Identification Strategies

The most common approach in applied research is to estimate equation (1) using instruments taken from the supply-side of the model (Berry and Haile, 2021; Gandhi and Nevo, 2021). For example, if $\alpha^{(k)} = 0$ and $\gamma^{(k)} \neq 0$, where $\alpha^{(k)}$ and $\gamma^{(k)}$ are the kth elements of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, then the characteristic $x^{(k)}$ is a cost-shifter that satisfies the exclusion restriction (in demand) and the relevance condition (in supply). Similarly, "markup-shifters" that create variation in $\mu_{jt}(\cdot)$ but do not enter equation (1) can be valid supply-side instruments. 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).

An important property of inverse demand and supply equations, equations (1) and (4), is that the slopes with respect to price are exact opposites. This provides a second path to identification: estimating the model through equation (4), using instruments taken from the demand-side of the model. For example, if $\alpha^{(k)} \neq 0$ and $\gamma^{(k)} = 0$, then $x^{(k)}$ is a demand-shifter that satisfies the exclusion restriction (in supply) and the relevance condition (in demand). Markup-shifters also can be valid instruments for this purpose. To our knowledge, the idea that exogenous demand-side variation identifies the parameters of the model—potentially without any exogenous supply-side variation—has not previously been recognized in the literature. Berry et al. (1995) use markup-shifters to jointly estimate the supply and demand equations, and the simultaneous equations framework helps clarify why this improves efficiency.

The third path to identification is to place a direct restriction on the relationship of the structural error terms, along the lines of $Cov(\xi_{jt}, \eta_{jt}) = 0$. In our baseline model, the assumption of uncorrelated demand and cost shocks implies constant marginal costs, though this can be relaxed (Section 3.4). Whether uncorrelatedness is credible depends in part on whether covariates or fixed effects absorb otherwise confounding variation (Section 3.5). Covariance restrictions have been studied previously in the context of perfect competition, where the slope of the supply equation is generated by increasing marginal cost functions (e.g., Hausman and Taylor, 1983; Hausman et al., 1987; Matzkin, 2016). We consider the case in which the slope of the supply equation may be affected market power, which we develop in the next section.

These three approaches all impose an orthogonality condition involving one or more structural error terms. There are nonetheless important distinctions. In particular, supply-side instruments allow for identification with an informal understanding of supply and with nonpara-

metric demand (Berry and Haile, 2014), whereas demand-side instruments and covariance restrictions require a formal supply-side model and parametric assumptions on demand. Thus, when valid supply-side instruments are available, the standard approach to estimation can proceed under weaker assumptions. Another distinction is that the instrument-based approaches require a relevance condition to be satisfied. In contrast, the covariance restriction exploits all of the endogenous variation in the data and there is no "first-stage" relevance condition that must be met; estimation can proceed even when valid instruments are unavailable or weak.

The covariance restrictions approach we develop is distinct from the existing results on identification of simultaneous equations with covariance restrictions. Hausman and Taylor (1983) and Hausman et al. (1987) outline a strategy that uses a covariance restriction to generate a residual instrument that can be employed in two-stage least-squares (2SLS) estimation. Importantly, consistent estimation in that context requires that a valid and relevant instrument exists, in addition to the covariance restriction, whereas we have no such requirement. Instead, we use employ economic theory to specify the relationship between supply and demand. We provide a formal exposition of the connection in Appendix B.

2.3   The Empirical Content of Supply-Side Restrictions

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

As we establish formally in the following section, the maintained assumptions about supply and demand yield a function that maps the endogenous parameter β to $Cov(\xi, \eta)$. Following this mapping, $Cov(\xi, \eta)$ can act as a free parameter that rationalizes any candidate value of β, conditional on the data and the structure of demand and supply. Thus, supply restrictions alone do not assist with identification in general. However, as we show in Section 3.3, there are special cases where the supply-side model can bound the possible range of $Cov(\xi, \eta)$ and thus rule out some values of β.

Efficiency gains are possible if additional moments are imposed along with the supply-side assumptions. One prominent example is Berry et al. (1995), which estimates both the demand and supply relations (1) and (4) with markup shifters as instrumental variables, combining the first two identification strategies described above. The efficiency gains from joint estimation can be attributed to the additional moment conditions. If marginal cost shifters are the only available instruments (e.g., Nevo, 2001) then imposing the supply-side does not yield efficiency improvements because cost shifters do not separately identify the supply equation.

Efficiency gains are also possible through restrictions on the covariance structure of unob-

servables. Some applications that use the "optimal instruments" method make the assumption of homoskedasticity in the unobserved demand and cost shocks, which can provide modest efficiency gains (Reynaert and Verboven, 2014; Conlon and Gortmaker, 2020).

3 Covariance Restrictions

3.1 Identification

We now formalize the identification argument under a covariance restriction between the unobserved demand and marginal cost shocks. Let β^{OLS} denote the probability limit of the OLS estimate of the price coefficient obtained from a regression of $h(\cdot)$ on p and $\boldsymbol{x}$, where we use the absence of subscripts to denote random variables. Then,

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

where p^* is the component of price orthogonal to $\boldsymbol{x}$, i.e., $p^* = p - \boldsymbol{x}\mathbb{E}[\boldsymbol{x}'\boldsymbol{x}]^{-1}\mathbb{E}[\boldsymbol{x}'p]$. The residuals of a regression of p on $\boldsymbol{x}$ provide approximate realizations of p^*.

Our main identification connects the price coefficient to the covariance of the structural error terms and empirical variation present in the data:

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

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

Therefore, β solves the following quadratic equation:

$$\begin{aligned}
0 = \ & \beta^2 \\
& + \left(\frac{Cov\left(p^*, \lambda\right)}{Var(p^*)} - \beta^{OLS} + \frac{Cov(\xi, \eta)}{Var(p^*)} \right) \beta \\
& + \left(-\beta^{OLS} \frac{Cov\left(p^*, \lambda\right)}{Var(p^*)} - \frac{Cov\left(\xi^{OLS}, \lambda\right)}{Var(p^*)} \right)
\end{aligned} \tag{7}$$

All proofs are in Appendix C. Aside from β and $Cov(\xi, \eta)$, all of the terms in equation (7) can be constructed from data. In particular, equation (7) links the probability limit of the reduced-form OLS coefficient to the true causal parameter.

There are at most two solutions for β for any given value of $Cov(\xi, \eta)$. Further, in most empirical models, β is likely to be the lower root of equation (7). Our second identification

result provides formal conditions under which this is the case:

Proposition 2. *The parameter β is the lower root of equation* (7) *if and only if*

$$-\frac{1}{\beta}Cov(\xi,\eta) \leq Cov\left(p^*, \eta - \frac{1}{\beta}\xi\right) \tag{8}$$

and, furthermore, β is the lower root of equation (7) *if*

$$0 \leq \beta^{OLS}Cov\left(p^*, \lambda\right) + Cov\left(\xi^{OLS}, \lambda\right). \tag{9}$$

In the first condition, it is helpful to think of $-\frac{1}{\beta}\xi$ as the demand-side structural error term, scaled so that units are equivalent to those of marginal costs (and price). If $Cov(\xi,\eta) = 0$, the condition holds as long as prices tend to increase with demand and marginal costs, as occurs in most empirical models. Thus, β is likely the lower root of equation (7) in most applications. The second condition is derived using properties of the quadratic formula. Because the terms in equation (9) are constructed from data, the sufficient condition may be estimated and used to test (and possibly reject) the null hypothesis that multiple negative roots exist. Henceforth, we assume that β is the lower root.

3.2 Two Approaches to Estimation

We have shown that the unknown price coefficient β is a function of $Cov(\xi,\eta)$. Because of this, we focus on covariance restrictions in first moments, though higher-order restrictions can work. We further restrict attention to uncorrelated demand and cost shocks. In the context of the baseline model, this assumption implies constant marginal costs, which is sometimes maintained in the literature (e.g., Miller and Weinberg, 2017; Backus et al., 2021). Estimation can employ the analytical solution to equation (7) or use the method-of-moments (e.g., Döpper et al., 2021). With the former approach, a three-stage estimator for β using the covariance restriction $Cov(\xi,\eta) = 0$ takes the form:

Corollary 1. *(Three-Stage Estimator) If β is the lower root of equation* (7) *and $Cov(\xi,\eta) = 0$, then a consistent estimate of β is given by*

$$\hat{\beta}^{\text{3-Stage}} = \frac{1}{2}\left(\hat{\beta}^{OLS} - \frac{Cov\left(p^*, \lambda\right)}{Var(p^*)} - \sqrt{\left(\hat{\beta}^{OLS} + \frac{Cov\left(p^*, \lambda\right)}{Var(p^*)}\right)^2 + 4\frac{Cov\left(\hat{\xi}^{OLS}, \lambda\right)}{Var(p^*)}}\right) \tag{10}$$

The estimator is consistent for the lower root of equation (7). It can be calculated in three stages: (i) regress h on p and x with OLS, (ii) regress p on x with OLS and obtain the residuals p^*, and (iii) construct the estimator as shown.

The method-of-moments approach recasts the covariance restriction as an orthogonality condition: $\mathbb{E}[\xi_{jt}\eta_{jt}] = 0$. The corresponding estimator is:

$$\hat{\beta}^{MM} = \arg\min_{\tilde{\beta}<0} \left[\frac{1}{T}\frac{1}{J}\sum_t\sum_{j\in\mathcal{J}}\xi_{jt}(\tilde{\beta};h_{jt},x_{jt},\boldsymbol{\alpha})\eta_{jt}(\tilde{\beta};\lambda_{jt},x_{jt},\boldsymbol{\gamma})\right]^2 \tag{11}$$

where $\xi(\tilde{\beta};h,x,\boldsymbol{\alpha})$ and $\eta(\tilde{\beta};\lambda,x,\boldsymbol{\gamma})$ are the estimated residuals given the data and the candidate parameter. Care must be taken to ensure convergence to the lower root. Still, this approach may be preferred to the three-stage estimator. First, additional moments can be incorporated in estimation, allowing for efficiency improvements and specification tests (e.g., Hausman, 1978; Hansen, 1982). Second, the three-stage estimator requires orthogonality between ξ and all regressors (i.e., $E[\boldsymbol{x}\xi] = \boldsymbol{0}$), whereas the numerical approach can be pursued under a weaker assumption that allows for correlation between ξ and covariates that enter the cost function only.

With either approach to estimation, the empirical variation that identifies β is the relative variance in quantity and price in the data. For the case in which $Cov(\xi,\eta) = 0$, we obtain a formal result:

Proposition 3. *If $Cov(\xi,\eta) = 0$, then a first-order approximation to lower root of equation* (7) *is*

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

where h^* is the component of h orthogonal to $\boldsymbol{x}$, i.e., $h^* = h - \boldsymbol{x}\mathbb{E}[\boldsymbol{x}'\boldsymbol{x}]^{-1}\mathbb{E}[\boldsymbol{x}'h]$. The residuals of a regression of h on $\boldsymbol{x}$ provide approximate realizations of h^*.

Intuition for this result can be gleaned from the simultaneous equations representation of the model (Section 2.2), in which β determines the slope of both demand and supply. A large β corresponds to a flatter inverse demand schedule (i.e., price sensitive consumers) and a flatter inverse supply schedule (i.e., less market power). Uncorrelated shifts in such schedules tend to generate more variation in quantity than price. By contrast, a small β corresponds to steeper inverse demand and inverse supply schedules, such that uncorrelated shifts generate more variation in price than quantity. Connecting these observations formally generates an approximation of the lower root based on the ratio of variances. Estimation with a covariance restriction converts the endogenous variation in prices and quantity into consistent estimates.

3.3 Analysis of Bounds

Our result in Proposition 1 above characterizes the model-implied constraints on the price parameter based on observed data. Even without exact knowledge of $Cov(\xi,\eta)$, it is possible

to use equation (7) to generate bounds on the price parameter. We first consider bounds that utilize prior knowledge of the sign of the correlation between demand and supply shocks. Next, we show how the model and the data together may bound the price coefficient without any additional information.

In settings where an assumption of uncorrelated demand and cost shocks is inappropriate, the researcher may expect the sign of the correlation to go in a particular direction (positive or negative). With a prior of the sign of $Cov(\xi, \eta)$, bounds can be placed on β. The reason is that there is a one-to-one mapping between the value of $Cov(\xi, \eta)$ and the lower root of equation (7):

Lemma 1. *(Monotonicity) Under assumptions 1 and 2, a valid lower root of equation* (7) *(i.e., one that is negative) is decreasing in $Cov(\xi, \eta)$. The range of the function is $(0, -\infty)$.*

Thus, if higher quality products are more expensive to produce ($Cov(\xi, \eta) \geq 0$) or firms invest to lower the marginal costs of their best-selling products ($Cov(\xi, \eta) \leq 0$), then one-sided bounds can be placed on β. More generally, let $r(m)$ be the lower root of the quadratic in equation (7), evaluated at $Cov(\xi, \eta) = m$. Then $Cov(\xi, \eta) \geq m$ produces $\beta \in (-\infty, r(m)]$, and $Cov(\xi, \eta) \leq m$ produces $\beta \in [r(m), 0)$. The lower root, $r(m)$, can be estimated with the method-of-moments.[4]

Even without any prior knowledge about $Cov(\xi, \eta)$, we can show that some values of the price parameter cannot rationalize the data. Thus, the demand and supply assumptions alone may be informative about the range of β. This result may be somewhat surprising. Formally, this occurs when the quadratic from equation (7) does not have a lower root, and thus no valid solution for β.

To see how this may occur, represent the quadratic from equation (7) as $az^2 + bz + c$, keeping in mind that one root is $\beta < 0$. 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 $Cov(\xi, \eta)$, which is a component of b. From the monotonicity result (Lemma 1), we can use the excluded values of $Cov(\xi, \eta)$ from this result to rule out values of β.

We now state the result formally:

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

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

The bound exists if and only if the term inside the radical is non-negative. Further, through equation (7), this lower bound on $Cov(\xi, \eta)$ provides an upper bound on β.

---

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

From the monotonicity result (Lemma 1), we can use the excluded values of $Cov(\xi, \eta)$ from this result to rule out values of β. A model-based upper bound for β is obtained by evaluating the lower root of equation (7) at the model-based bound of $Cov(\xi, \eta)$.

In practice, priors over the the covariance of unobserved shocks may be combined with model-based bounds to further restrict the identified set. We explore these approaches further in an application in Section 5.3.

3.4   Extensions to the Baseline Model

The method-of-moments estimator developed above—which relies on uncorrelatedness between demand and cost shocks—can readily extend to other environments beyond that of our baseline model. Consider a setting in which marginal costs vary with output due to underlying production technology. For example:

$$mc_{jt}(\chi_t; \boldsymbol{\theta}) = \boldsymbol{x}_{jt}'\boldsymbol{\gamma} + g(q_{jt}; \tau) + \eta_{jt} \tag{13}$$

where $g(q_{jt}; \tau)$ is a potentially nonlinear function that explicitly incorporates the relationship between output and marginal cost. The inverse supply equation becomes:

$$\lambda_{jt}(\boldsymbol{q}_t, \boldsymbol{p}_t, D(\chi_t); \boldsymbol{\theta}) = -\beta p_{jt} + \beta \boldsymbol{x}_{jt}'\boldsymbol{\gamma} + \beta g(q_{jt}; \tau) + \beta \eta_{jt} \tag{14}$$

In the augmented model, both markup adjustments and non-constant marginal costs can contribute to price endogeneity, even if $Cov(\xi, \eta) = 0$.

The OLS regression of $h(q_{jt}, w_{jt}; \sigma)$ on p and x yields a price coefficient with the following probability limit:

$$\beta^{OLS} = \beta - \frac{1}{\beta} \frac{Cov\,(\xi, \lambda)}{Var(p^*)} + \frac{Cov(\xi, g(q; \tau))}{Var(p^*)} + \frac{Cov(\xi, \eta)}{Var(p^*)} \tag{15}$$

As in equation (6), it is possible to expresses the probability limit in terms of OLS residuals, ξ^{OLS}, instead of ξ. The above expression explicitly decomposes the bias due to marginal costs into components based on a known function (g) and the residuals (η). Without prior knowledge of τ, an additional restriction is necessary to pin down $g(q_{jt}; \tau)$ and extend the identification results of the preceding sections. For any candidate $\hat{\tau}$, however, a corresponding estimate of $\beta(\hat{\tau})$ can be obtained:

Proposition 5. *When marginal costs take the semi-linear form of equation* (13) *and residual*

demand and cost shocks are uncorrelated, β solves the following quadratic equation:

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

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

Proof. See appendix.

Thus, it is possible to control for a direct relationship between output and marginal costs with some additional structure. After controlling for g, it may be reasonable to assume that the residual cost shocks (η) are uncorrelated with demand. Under the assumption that $Cov(\xi, \eta) = 0$, additional covariance restrictions or other instruments may be employed to pin down τ and thus identify β.

In general, the framework introduced in the preceding sections admits extensions that are tailored to different economic environments. Our analysis shows how a single covariance restriction plays a specific and important role, allowing the price coefficient to be identified when firms adjust prices in response to marginal costs. Other sources of endogeneity, such as marginal costs that increase with output, require the use of additional restrictions. As shown above, restrictions that identify the cost function do not, by themselves, resolve the source of price endogeneity that is central in our analysis. However, it is straightforward to incorporate both types in estimation.

3.5 Assessment

The covariance restrictions approach to estimation can be both useful and credible when exogenous variation in marginal costs exists, yet the cost-shifters that give rise to the variation are unobserved by the econometrician. As an illustrative example, Döpper et al. (2021) estimate demand for consumer products using covariance restrictions. Given the specification of the marginal cost function, the structural error includes two sources of variation that have been exploited as instruments in recent research: product-specific changes in input costs (Backus et al., 2021) and product-specific changes in distribution costs (Miller and Weinberg, 2017). Döpper et al. (2021) employ a rich set of fixed effects to better isolate exogenous variation in costs, and they obtain demand elasticities that are similar to those reported in the literature.

Credibility also depends on whether the supply-side structural error term contains additional, confounding variation. In some applications, it can be possible to absorb the most

obvious sources of confounding variation using fixed effects or other controls. To make this explicit, suppose that the demand and cost functions are given by:

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

where, again, the subscripts j and t refer to products and markets, respectively. The unobserved error terms are $\xi_{jt} = D_j + F_t + E_{jt}$ and $\eta_{jt} = U_j + V_t + W_{jt}$. If products with higher quality have higher marginal costs then $Cov(U_j, D_j) > 0$. The econometrician can account for the relationship by estimating D_j for each firm; the residual $\xi^*_{jt} = \xi_{jt} - D_j$ is uncorrelated with U_j. Similarly, if costs are higher (or lower) in markets with high demand then $Cov(F_t, V_t) \neq 0$, but market fixed effects can be incorporated to absorb the confounding variation. Thus, it can be possible to isolate components of the error terms over which a covariance restriction is credible.

4   Small-Sample Performance

We use Monte Carlo simulations to examine the small sample performance of the three estimation strategies discussed in Section 2.2. First, we consider the impact of relative variation in demand and supply shocks on the performance of the estimators. Second, we consider a misspecification of the supply-side model in which the econometrician makes an assumption about conduct that does not match the data-generating process. By making use of both demand-side and cost-side variation, the covariance restrictions approach can mitigate to some extent potential sources of bias.

4.1   Relative Variation in Cost and Demand Shocks

The first set of simulations consider the empirical performance of different identification strategies and the relation to the underlying variation in demand and supply shocks. We examine a monopolist with a marginal cost of $mc_t = \eta_t$ and a linear demand schedule of $q_t = 10 - p_t + \xi_t$. Thus, $\beta = -1$. We let ξ_t and η_t have independent uniform distributions. We consider four specifications: (i) $\xi \sim U(0, 2)$ and $\eta \sim U(0, 8)$, (ii) $\xi \sim U(0, 4)$ and $\eta \sim U(0, 6)$, (iii) $\xi \sim U(0, 6)$ and $\eta \sim U(0, 4)$, and (iv) $\xi \sim U(0, 8)$ and $\eta \sim U(0, 2)$. Moving from (i) to (iv), demand-side variation increases and supply-side variation decreases.

As is well known, if both cost and demand variation is present, then equilibrium outcomes provide a "cloud" of data points that do not necessarily correspond to the demand curve. To illustrate, we present one simulation of 500 observations from each specification in Figure 1, along with the fit of an OLS regression of quantity on price. The probability limits of the OLS estimate in each scenario are -0.882, -0.385, 0.385, and 0.882. With greater demand-

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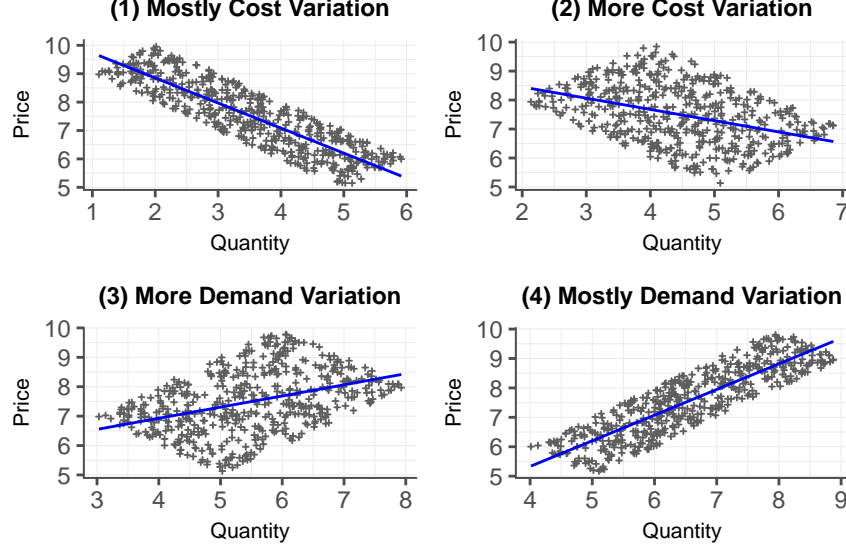
Notes: This figure displays equilibrium prices and quantities under four different specifications for the distribution of unobserved shocks to demand and marginal costs. The line in each figure indicates the slope obtained by OLS regression.

side variation, the endogeneity bias is larger.[5] Any of the three identification strategies can be used to resolve endogeneity bias in this setting. However, in small samples, it is known that instrumental variables may suffer from bias and have a large variance.

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 assume $Cov(\xi, \eta) = 0$. The estimate of β is simply $-\sqrt{Var(q)/Var(p)}$ because the approximation of equation (12) is exact in this context. For a supply-side instrument, we estimate demand with 2SLS using the cost shock η_t as the instrument. That is, we use all of the supply-side variation to estimate demand. For the demand-side instrument, we estimate the supply equation with 2SLS using ξ_t as the instrument. Note that all three approaches rely on an identical orthogonality condition: $\mathbb{E}[\xi_t \cdot \eta_t] = 0$.

Table 1 summarizes the results.[6] 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

---

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

[6]To avoid outliers arising from the weak instrument problem, we bound the estimates of β on the range $[-100, 100]$. For specifications that suffer from weak instruments, this biases the standard errors toward zero. This affects specifications where the estimated standard error is greater than one, i.e., in 9 of 48 specifications.

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	-1.004	(0.098)	-1.017	(0.201)	-1.018	(0.206)	-1.005	(0.099)
50	-1.001	(0.068)	-1.008	(0.136)	-1.007	(0.135)	-1.001	(0.068)
100	-1.001	(0.047)	-1.003	(0.094)	-1.004	(0.093)	-1.001	(0.047)
500	-1.000	(0.021)	-1.001	(0.041)	-1.001	(0.042)	-1.000	(0.021)

(b) Supply Shifters (IV-1)

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

(c) Demand Shifters (IV-2)

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

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

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 2.6, 4.2, 7.3, and 32.6 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 also is due to a weak instruments problem. Thus, in settings where instruments perform poorly, a covariance restriction may still provide a precise estimate because it exploits all of the endogenous price and quantity variation in the data. In our simulations, the covariance restriction approach combines both sources of variation used by the instrument-based approaches. Moreover, it does not require that these sources of variation be observed.

4.2 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

Table 2: Small-Sample Properties: Supply-Side Misspecification

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

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

not match the data-generating process.[7] 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

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

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

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

where $\frac{\partial q}{\partial p}$ is a matrix of demand derivatives and $\circ$ denotes element-by-element multiplication. The model nests Bertrand competition ($\kappa = 0$) and joint price-setting behavior ($\kappa = 1$), as well as capturing (non-micro-founded) intermediate cases.

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 ($\kappa = 0$),

---

[7]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.

thus generating supply-side misspecification.

Table 2 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 receives some protection against bias from model misspecification.

5   Empirical Applications

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.[8]

Let the indirect utility that consumer i receives from product j in market t (a combination

---

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

of a time period and a region) be given by

$$u_{ijt} \quad = \quad \delta_{jt}(p_{jt};\alpha) + \mu_{ijt}(p_{jt}, \boldsymbol{D}_i, \boldsymbol{v}_i; \boldsymbol{\Pi}, \boldsymbol{\Sigma}) + \epsilon_{ijt} \tag{16}$$

where the mean utility of each product, $\delta_{jt}(\cdot)$, and contribution of demographics to consumer-specific deviations, $\mu_{ijt}(\cdot)$, respectively are given by

$$
\begin{aligned}
\delta_{jt}(p_{jt};\beta,\boldsymbol{\alpha}) &= \beta p_{jt} + \boldsymbol{x}'_{jt}\boldsymbol{\alpha} + \xi_j + \Delta\xi_{jt} \\
\mu_{ijt}(p_{jt}, \boldsymbol{D}_i, \boldsymbol{v}_i; \boldsymbol{\Pi}, \boldsymbol{\Sigma}) &= [p_{jt}, \boldsymbol{x}_{jt}]' * (\boldsymbol{\Pi}\boldsymbol{D}_i + \boldsymbol{\Sigma}\boldsymbol{v}_i)
\end{aligned}
$$

with $\boldsymbol{D}_i$ and $\boldsymbol{v}_i$ being vectors of consumer-specific demographic characteristics.

The probability with which consumer i selects product j is

$$s_{ijt}(\boldsymbol{\delta}_t, p_{jt}, \boldsymbol{D}_i, \boldsymbol{v}_i; \boldsymbol{\Pi}, \boldsymbol{\Sigma}) = \frac{\exp(\delta_{jt}(p_{jt};\beta,\boldsymbol{\alpha}) + \mu_{ijt}(p_{jt}, \boldsymbol{D}_i, \boldsymbol{v}_i; \boldsymbol{\Pi}, \boldsymbol{\Sigma}))}{1 + \sum_{k=1}^{J_{kt}} \exp(\delta_{kt}(p_{kt};\beta,\boldsymbol{\alpha}) + \mu_{ikt}(p_{kt}, \boldsymbol{D}_i, \boldsymbol{v}_i; \boldsymbol{\Pi}, \boldsymbol{\Sigma}))} \tag{17}$$

where $\boldsymbol{\delta}_t = (\delta_{1t}, \delta_{2t}, \dots)$ is the vector of mean utilities. The market share of product j is obtained by integrating over the joint distribution of consumer demographics:

$$s_{jt}(\boldsymbol{\delta}_t, p_{jt}; \boldsymbol{\Pi}, \boldsymbol{\Sigma}) = \frac{1}{I} \sum_i s_{ijt}(\boldsymbol{\delta}_t, p_{jt}, \boldsymbol{D}_i, \boldsymbol{v}_i; \boldsymbol{\Pi}, \boldsymbol{\Sigma}) \tag{18}$$

We now connect to the modeling framework of Section 2. Let the set of products sold by the same firm as product j be given by $\mathcal{J}_{f(j)}$. Then, under the assumption of Bertrand price competition, we have:

$$\lambda_{jt}(\boldsymbol{s}_t, \boldsymbol{p}_t; \boldsymbol{\theta}) = \frac{s_{jt}}{\frac{1}{I}\sum_i s_{ijt}(1 - s_{ijt})} - \sum_{k \in \mathcal{J}_{f(j)}\backslash j} \frac{s_{kt}}{\frac{1}{I}\sum_i s_{ijt}s_{ikt}} \tag{19}$$

where the denominators integrate over the (product of) consumer-specific choice probabilities. From an econometric standpoint, λ_{jt} is free from the price parameter β because it depends only on market shares and consumer-specific choice probabilities. The market shares are data. From equation (17), the consumer-specific choice probabilities depend on $\mu_{ijt}(\cdot)$, which obtains immediately from data and $(\boldsymbol{\Pi}, \boldsymbol{\Sigma})$, and on $\boldsymbol{\delta}_t(\cdot)$, which obtains from the contraction mapping of Berry et al. (1995), again given data and $(\boldsymbol{\Pi}, \boldsymbol{\Sigma})$.

Marginal costs in the Nevo (2000) model are given by

$$mc_{jmt} \quad = \quad \eta_j + \Delta\eta_{jt} \tag{20}$$

where η_j is a product fixed effect, and $\Delta\eta_{jt}$ is the supply-side structural error term.

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 is sometimes 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 $(\beta, \boldsymbol{\alpha})$, as in our baseline model from Section 2, and also $(\boldsymbol{\Pi}, \boldsymbol{\Sigma})$. Therefore, additional identifying assumptions are needed. Some recent applications use so-called "micro-moments" constructed from the observed behavior of individual consumers (e.g., Backus et al., 2021; Döpper et al., 2021) or "second-choice" data on what consumers view as their next-best option (e.g., Grieco et al., 2021). Both of these strategies identify $(\boldsymbol{\Pi}, \boldsymbol{\Sigma})$ but not the price parameter (Berry and Haile, 2020), which supports 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 3 summarizes the results of estimation based on the instruments (panel (a)) and covariance restrictions (panel (b)). Both identification strategies yield similar mean own-price demand elasticities: -3.70 with instruments and -3.61 with covariance restrictions. Overall, the different approaches produce similar patterns for the coefficients. Most of the point estimates under covariance restrictions fall in the 95 percent confidence intervals implied by the specification with instruments, including that of the mean price parameter. 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. 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. Given the marginal cost specification in equation (2), the baseline model accommodates

Table 3: 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.

increasing marginal costs only if $Cov(\xi, \eta) > 0$. However, if the upward-sloping component of marginal costs can be modeled more explicitly, then the covariance restriction approach to estimation may remain valid. To illustrate, we consider the setting and data of Fowlie et al. (2016) ["FRR"], which examines market power in the cement industry.

The model features Cournot competition among cement plants facing capacity constraints. Let the market demand curve in region r and year t have a logit form:

$$h_{rt}(Q_{rt}) \equiv \ln(Q_{rt}) - \ln(M_r - Q_{rt}) = \boldsymbol{x}'_{rt}\boldsymbol{\alpha} + \beta p_{rt} + \xi_{rt} \tag{21}$$

where $Q_{rt} = \sum_{j \in \mathcal{J}} q_{jrt}$ is total quantity and M_r is the "market size" of the region.[9] Further, we allow marginal costs to vary with quantity according to

$$mc_{jrt} = \boldsymbol{x}'_{jrt}\boldsymbol{\gamma} + g_{jrt}(q_{jrt}) + \eta_{jrt} \tag{22}$$

---

[9]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 $M_r = 2 \times \max_t\{Q_{rt}\}$.

In particular, we follow FRR and assume that that g is a "hockey stick" function, $g_{jrt}(q_{jrt}) = 2\psi 1\{q_{jrt}/k_{jr} > 0.9\}(q_{jrt}/k_{jr} - 0.9)$, where k_{jr} and q_{jrt}/k_{jr} are capacity and utilization, respectively. Marginal costs are constant if utilization is less than 90%, and increasing linearly at rate determined by $\psi \geq 0$ otherwise.

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 $Cov(\xi_{rt}, \bar{\eta}_{rt}) = 0$, where $\bar{\eta}_{rt} = \frac{1}{J}\eta_{jrt}$. 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 (4) readily admits Cournot competition. As only market-level price and costs measures are observed, one must use the mean firm-level quantity $\bar{q}_{rt} = \frac{1}{J}Q_{rt}$ 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 3.4 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). We solve the quadratic equation expressed in Proposition 5 to recover $\hat{\beta}$, under the assumption that residual demand and cost shocks are uncorrelated. Here, whether the covariance restriction is reasonable depends primarily on the relationship between construction activity (unobserved demand) and the prices of coal and electricity (unobserved supply costs). In this context, there is a theoretical basis for 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.

We find that the covariance restrictions approach yields a demand elasticity of -1.15, with a standard error of 0.18.[10] This is nearly identical 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. This similarity reflects, we believe, that the identifying assumptions are actually quite 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 (standard error of 0.15) using OLS. And, if we use the covariance restriction without accounting for the presence of $g_{jrt}(\cdot)$, we obtain a demand elasticity of -0.90 (standard error of 0.13), which is in between the OLS and 2SLS estimates.

---

[10]There are 520 region-year observations over 1984-2009. The demand specification includes region fixed effects. We obtain bootstrapped standard errors based on 200 random samples constructed by drawing from the data with replacement.

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"].[11] The economics of the industry suggest that the covariance restriction $Cov(\xi, \eta) = 0$ 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, 2021). Thus, all else equal, greater demand generates more higher marginal costs, inclusive of the opportunity cost. Absent a model of these opportunity costs, it is difficult to achieve point identification using the covariance restriction. Instead, we use the industry to illustrate the bounds approach to identification, with multiple sets of bounds.

The AH model features differentiated-products Bertrand competition among firms facing a nested logit demand system. Products are classified into the following groups: nonstop flights, one-stop flights, and the outside good. The nested logit demand system can be expressed as

$$\ln s_{jmt} - \ln s_{0mt} - \sigma \ln \overline{s}_{jmt|g} = \beta p_{jmt} + x'_{jmt}\alpha + \xi_{jmt} \tag{23}$$

where s_{jmt} is the market share of product j in market m in period t. The conditional market share, $\overline{s}_{j|g} = s_j / \sum_{k \in g} s_k$, is the the choice probability of product j given that its "group" of products, g, is selected. The outside good is indexed as $j = 0$. Higher values of σ increase within-group consumer substitution relative to across-group substitution.[12]

Given the role of opportunity costs in the industry, we assume $Cov(\xi_{jmt}, \eta_{jmt}) \geq 0$. Under that assumption, we reject values (β, σ) that produce a negative correlation in product-specific shocks. We combine this with model-based bounds (Section 3.3). Finally, if the correlation in product-level shocks is weakly positive, it is reasonable to assume that the group-level shocks are also weakly positive, through a similar deduction. Thus, we apply the group-level inequality

$$E_{gmt}[\overline{\xi}_{gmt} \cdot \overline{\eta}_{gmt}] \geq 0, \tag{24}$$

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

Figure 2 displays the rejected regions based on the model and our assumptions on unobserved shocks. The gray region corresponds to the parameter values rejected by the model-based bounds; the model itself rejects some values of β if $\sigma \geq 0.62$. As σ becomes larger, a

---

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

[12]The covariates include an indicator for nonstop itineraries, the distance between the origin and destination cities, and a measure of the airline's "hub sizes" at the origin and destination cities. We also include airline fixed effects and route×quarter fixed effects. Market size, which determines the market share of the outside good, is equal to the total population in the origin and destination cities.

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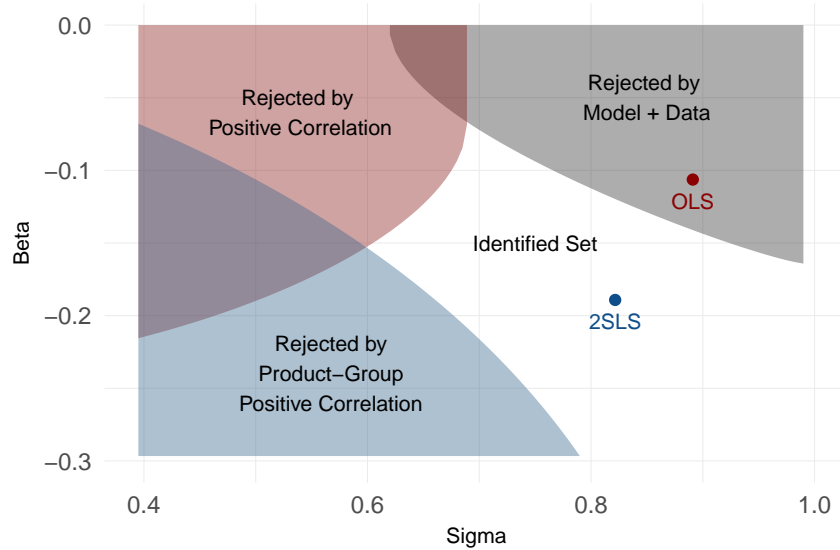
Notes: This figure displays candidate parameter values for (σ, β). The gray region indicates the set of parameters that cannot generate the observed data from the assumptions of the model. The red region indicates the set of parameters that generate $Cov(\xi, \eta) < 0$, and the blue region indicates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

more negative β is required to rationalize the data. The dark red region corresponds to parameter values that generate negative correlation between demand and supply shocks. The region is rejected under the prior that $Cov(\xi_{jmt}, \eta_{jmt}) \geq 0$. The dark blue region provides the corresponding set for the prior $Cov(\overline{\xi}_{gmt}, \overline{\eta}_{gmt}) \geq 0$ 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 β, as these lower values cannot generate positive correlation in both product-level and product-group-level shocks. Similarly, we obtain an upper bound on β of -0.067 across all values of σ. For context, we plot the OLS and the 2SLS estimates in Figure 2. The OLS estimate falls in a rejected region and can be ruled out by the model alone. The 2SLS estimate falls within the identified set. This result is not mechanical, as these point estimates are generated with non-nested assumptions.

6   Conclusion

Our objective has been to evaluate the identifying power of supply-side assumptions in models of imperfect competition. Invoking the supply model in estimation expands the set of restrictions that obtain identification. In particular, many applications in applied microeconomics

currently employ instruments from the supply-side of the model in estimation. We show that demand-side instruments and covariance restrictions between unobserved demand and cost shocks also can resolve price endogeneity and allow for consistent estimation. The covariance restrictions approach is notable in part because there is no relevance condition; instead the endogenous variation in quantity and price is interpreted through the lens of the model to recover the structural parameters. As this is somewhat novel, we provide three empirical applications to demonstrate how covariance restrictions can be applied and evaluated.

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

References

AGUIRREGABIRIA, V. AND C.-Y. HO (2012): "A Dynamic Oligopoly Game of the US Airline Industry: Estimation and Policy Experiments," *Journal of Econometrics*, 168, 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.

——— (2021): "Foundations of Demand Estimation," In Ho, K., Hortacsu, A., and Lizzeri, A. (Eds.), *Handbook of Industrial Organization, 4th Edition*. Elsevier.

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.

CILIBERTO, F., C. MURRY, AND E. TAMER (2021): "Market Structure and Competition in Airline Markets," *Journal of Political Economy*, 129, 2995–3038.

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.

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 (2021): "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," *The 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 (2021): "The Evolution of Market Power in the US Automobile Industry," Working Paper.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

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.

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 (2021): "Repositioning and market power after airline mergers," *RAND Journal of Economics*, forthcoming.

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

MILLER, N. H. AND M. WEINBERG (2017): "Understanding the Prices Effects of the MillerCoors Joint Venture," *Econometrica*, 85, 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 (2021): "Bounds on a Slope from Size Restrictions on Economic Shocks," Working paper.

REYNAERT, M. AND F. VERBOVEN (2014): "Improving the Performance of Random Coefficients Demand Models: The Role of Optimal Instruments," *Journal of Econometrics*, 179, 83–98.

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.

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.

WEGGE, L. (1965): "Identifiability Criteria for a System of Equations as a Whole," *The Australian Journal of Statistics*, 7, 67–77.

WILLIAMS, K. R. (2021): "The Welfare Effects of Dynamic Pricing: Evidence from Airline Markets," *Econometrica*, forthcoming.

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

For Online Publication

A   Demand System Applications

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

$$h_{jt}(\boldsymbol{q}_t, \boldsymbol{p}_t, \boldsymbol{x}_t, \boldsymbol{\xi}_t; \boldsymbol{\theta}) = \beta f(p_{jt}) + \boldsymbol{x}'_{jt}\boldsymbol{\alpha} + \xi_{jt} \tag{A.1}$$

Under this modified assumption, it is possible to employ a method-of-moments approach to estimate the structural parameters. When $f(p_{jt}) = \ln p_{jt}$, it is straightforward to extend our analytical identification results, under the modified assumptions that ξ is orthogonal to $\ln X$ and that $\ln \eta$ and ξ are uncorrelated. To see this, note that the optimal price for these demand systems takes the form $p_{jt} = \mu_{jt}c_{jt}$, where μ_{jt} is a markup that reflects demand parameters and (in general) demand shocks. It follows that the probability limit of an OLS regression of h on $\ln p$ is given by:

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

Therefore, the results developed in this paper 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

$$h_{jt}(\boldsymbol{q}_t, \boldsymbol{p}_t, \boldsymbol{x}_t, \boldsymbol{\xi}_t; \boldsymbol{\theta}) \equiv \ln(s_j) - \ln(s_0) - \sigma \ln(\overline{s}_{j|g})$$

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. The parameter $\sigma \in [0,1)$ determines the extent to which consumers substitute disproportionately among firms within the same group. If $\sigma = 0$ then the logit model obtains. We can construct the markup by calculating the total derivative of h with respect to s. For single-product firms at the Bertrand-Nash equilibrium,

$$\lambda_{jt} = \frac{dh_{jt}}{ds_{jt}}s_{jt} = \frac{1}{\frac{1}{1-\sigma} - s_{jt} - \frac{\sigma}{1-\sigma}\overline{s}_{j|g,t}}.$$

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

Building on the notation of Berry (1994) and Nevo (2000), let the indirect utility that consumer $i = 1, \ldots, I$ receives from product j be

$$
\begin{aligned}
u_{ij} =& \beta p_j + \boldsymbol{x}_j' \boldsymbol{\alpha} + \xi_j + p_j(\sigma_0 \zeta_i^0 + \pi_{01} D_{i1} + \ldots + \pi_{0d} D_{id}) \\
& + \sum_k x_j^k (\sigma_k \zeta_i^k + \pi_{k1} D_{i1} + \ldots + \pi_{kd} D_{id}) + \epsilon_{ij}
\end{aligned}
$$

where x_k^k is the kth element of x_j, ζ_i^k captures mean-zero consumer-specific tastes for characteristic k (including price), D_i captures consumer-specific mean-zero demographic characteristics, and ϵ_{ij} is a logit error. We have suppressed market subscripts for notational simplicity. Decomposing the right-hand side of the indirect utility equation into $\delta_j = \beta p_j + \boldsymbol{x}_j' \boldsymbol{\alpha} + \xi_j$ and $\phi_{ij} = p_j(\sigma_0 \zeta_i^0 + \pi_{01} D_{i1} + \ldots + \pi_{0d} D_{id}) + \sum_k x_j^k (\sigma_k \zeta_i^k + \pi_{k1} D_{i1} + \ldots + \pi_{kd} D_{id})$, the probability that consumer i selects product j is given by the standard logit formula

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

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

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

Thus, with the uncorrelatedness assumption the linear parameters can be recovered given the candidate parameter vector $\tilde{\sigma}$. We demonstrate how to estimate these parameters using additional covariance restrictions in our first application (Section 5.1), which also incorporates multi-product firms.

A.3 Constant Elasticity Demand

A special case that is often estimated in empirical work is when h and $f(p)$ are logarithms. With the modified demand assumption of equation (A.1), the constant elasticity of substitution (CES) demand model of Dixit and Stiglitz (1977) can be incorporated:

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

where q_t is an observed demand shifter, Π_t is a price index, and β provides the constant elasticity of demand. This model is often used in empirical research on international trade and firm productivity (e.g., De Loecker, 2011; Doraszelski and Jaumandreeu, 2013). Due to the constant elasticity, profit-maximization and uncorrelatedness imply $Cov(p, \xi) = 0$, and OLS produces

unbiased estimates of the demand parameters.[13] 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, $q_{jt} = \alpha_j + \sum_k \beta_{jk} p_k + \xi_{jt}$, which sometimes appears in identification proofs (e.g., Nevo, 1998) but is seldom applied empirically due to the large number of price coefficients. In principle, the system could be formulated such that $h(q_{jt}, w_{jt}; \sigma) \equiv q_{jt} - \sum_{k \neq j} \beta_{jk} p_k$. In addition to the own-product uncorrelatedness restrictions that could identify β_{jj}, one could impose cross-product covariance restrictions to identify β_{jk} ($j \neq k$). We discuss these cross-product covariance restrictions in our 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

We illustrate how our framework incorporates multi-product firms with the case of Bertrand pricing. Let K^m denote the set of products owned by multi-product firm m. When the firm sets prices on each of its products to maximize joint profits, there are $|K^m|$ first-order conditions, which can be expressed as

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

The market subscript, t, is omitted to simplify notation. For demand systems satisfying equation (1), $\frac{\partial q_k}{\partial p_j} = \beta \frac{1}{\frac{dh_j}{dq_k}}$, where the derivative $\frac{dh_j}{dq_k}$ is calculated holding the prices of other products fixed. Therefore, the set of first-order conditions can be written as

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

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

$$p_j = mc_j - \frac{1}{\beta} \left(\Lambda^m \boldsymbol{q}^m \right)_{i(j)}. \tag{A.3}$$

---

[13]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.

Here, $(\Lambda^m \boldsymbol{q}^m)_{i(j)}$ provides the entry corresponding to product j in the vector $\Lambda^m \boldsymbol{q}^m$. As the matrix Λ^m is not a function of the price parameter after conditioning on observables, this form of the first-order condition allows us to solve for β. Letting $\lambda \equiv (\Lambda^m \boldsymbol{q}^m)_{i(j)}$, we see that multi-product Bertrand fits in the class of models specified by equation (3).

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:

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

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

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

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

B Relationship to the Simultaneous Equations Literature

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

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

This system is analogous to equations (1) and (4) in the present paper. The two key differences are: by assumption, the covariate z_1 is excluded from the second equation, and there are two "slope" parameters (β_{12} and β_{21}). The linearity of the system is less consequential; for example, linearity also obtains in our setting with monopoly and linear demands.

The coefficients β_{12}, β_{21}, and γ_{11} are identified by invoking exogeneity of z_1, the exclusion restriction, and a covariance restriction: $E[z_1 \cdot \epsilon_1] = 0$, $E[z_1 \cdot \epsilon_2] = 0$, and $Cov(\epsilon_1, \epsilon_2) = 0$.[14] The parameters can be estimated jointly with GMM. Alternatively, the demand equation can be

---

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

estimated with 2SLS using z_1 as an instrument, and then the supply equation can be estimated with 2SLS using z_1 and the residual $\hat{\epsilon}_2 = y_2 - \hat{\beta}_{21y_1}$. Thus, in the presence of an initial valid instrument, a covariance restriction can be used to generate a residual instrument to employ in a standard instrumental variable estimator.[15] Importantly, consistent estimation requires that a valid and relevant instrument (z_1) exists.

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

C   Proofs

C.1   A Consistent and Unbiased Estimate for ξ

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

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

For some intuition, note that we can construct both the true demand shock and the OLS residuals as:

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

where this holds even in small samples. Without loss of generality, we assume $E[\xi] = 0$. The true demand shock is given by $\xi_0 = \xi^{OLS} + (\beta^{OLS} - \beta)p + \boldsymbol{x}'(\boldsymbol{\alpha}^{OLS} - \boldsymbol{\alpha})$. We desire to show that an alternative estimate of the demand shock, $\xi_1 = \xi^{OLS} + (\beta^{OLS} - \beta)p^*$, is consistent and unbiased. (This eliminates the need to estimate the true α parameters). It suffices to show that $(\hat{\beta}^{OLS} - \beta)p^* = (\hat{\beta}^{OLS} - \beta)p + \boldsymbol{x}'(\hat{\boldsymbol{\alpha}}^{OLS} - \boldsymbol{\alpha}) + \Upsilon$, where Υ is such that $E[\Upsilon = 0]$ and $\Upsilon \to 0$ as N gets large. It is straightforward to show this using the projection matrices for p and x.[16]

---

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

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

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

[16] Please contact the authors for the full proof.

C.2   Proof of Proposition 1 (Set Identification)

From equation (5), we have $\hat{\beta}^{OLS} \xrightarrow{p} \beta + \frac{Cov(p^*,\xi)}{Var(p^*)}$. The general form for a firm's first-order condition is $p = mc + \mu$, 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, X. By assumption, $c = \boldsymbol{x}'\boldsymbol{\gamma} + \eta$. If we substitute the first-order condition $p^* = \boldsymbol{x}'\boldsymbol{\gamma} + \eta + \mu - \hat{p}$ into the bias term from the OLS regression, we obtain

$$
\begin{aligned}
\beta^{OLS} - \beta = \frac{Cov(p^*,\xi)}{Var(p^*)} &= \frac{Cov(\xi, \boldsymbol{x}'\boldsymbol{\gamma} + \eta + \mu - \hat{p})}{Var(p^*)} \\
&= \frac{Cov(\xi,\eta)}{Var(p^*)} + \frac{Cov(\xi,\mu)}{Var(p^*)}
\end{aligned}
\tag{C.1}
$$

where the second line follows from the exogeneity assumption ($E[X\xi] = 0$).

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

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

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

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

Plugging this into equation (C.1) yields

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

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

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

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

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

C.3 Proof of Proposition 2 (Point Identification)

Part (1). We first prove the sufficient condition, i.e., that under assumptions 1 and 2, β is the lower root of equation (7) if the following condition holds:

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

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 (7), $a = 1$, and $4ac < 0$ if and only if equation (C.2) holds. Because the upper root is positive, $\beta < 0$ must be the lower root, and point identification is achieved given knowledge of $Cov(\xi,\eta)$. QED.

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

Lemma C.2. *The roots of equation (7) are β and $\frac{Cov(p^*,\xi)}{Var(p^*)} - \frac{Cov(p^*,\lambda)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}$.*

Proof of Lemma C.2. We first provide equation (7) for reference:

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

To find the roots, begin by applying the quadratic formula

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

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

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

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

$$p^* = \boldsymbol{x}'\boldsymbol{\gamma} + \eta + \mu - \hat{p}$$
$$= \boldsymbol{x}'\boldsymbol{\gamma} + \eta - \frac{1}{\beta}\lambda - \hat{p}$$

Therefore

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

and

$$Cov(\xi, \lambda) = -\beta\left(Cov(p^*, \xi) - Cov(\xi, \eta)\right)$$
$$\frac{Cov(\xi, \lambda)}{Var(p^*)} = -\beta\left(\frac{Cov(p^*, \xi)}{Var(p^*)} - \frac{Cov(\xi, \eta)}{Var(p^*)}\right) \tag{C.5}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$+ 2\beta\frac{Cov(\xi,\eta)}{Var(p^*)} - 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\xi)}{Var(p^*)} + 2\frac{Cov(\xi,\eta)}{Var(p^*)}\frac{Cov(p^*,\lambda)}{Var(p^*)}$$

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

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

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

The roots are given by

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

$$= \beta$$

and

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

which completes the proof of the intermediate result. QED.

Part (3). Consider the roots of equation (7), β and $\frac{Cov(p^*,\xi)}{Var(p^*)} - \frac{Cov(p^*,\lambda)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}$. The price parameter β may or may not be the lower root.[17] However, β is the lower root iff

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

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

$$\beta \;<\; -\beta\frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} + \beta\frac{Cov(p^*,p^*-c)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}$$

$$\beta \;<\; \beta\frac{Var(p^*)}{Var(p^*)} - \beta\frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} - \beta\frac{Cov(p^*,\eta)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}$$

$$0 \;<\; -\beta\frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} - \beta\frac{Cov(p^*,\eta)}{Var(p^*)} - \frac{Cov(\xi,\eta)}{Var(p^*)}$$

$$0 \;<\; \frac{Cov(p^*,-\frac{1}{\beta}\xi)}{Var(p^*)} + \frac{Cov(p^*,\eta)}{Var(p^*)} + \frac{1}{\beta}\frac{Cov(\xi,\eta)}{Var(p^*)}$$

The third line relies on the expression for the markup, $p - c = -\frac{1}{\beta}\frac{dh}{dq}q$. The final line holds because $\beta < 0$ so $-\beta > 0$. It follows that β is the lower root of equation (7) iff

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

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

---

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

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

because, in that case,

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

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

C.4 Proof of Proposition 3 (Approximation)

The demand and supply equations are given by:

$$h = \beta p + \boldsymbol{x}' \boldsymbol{\alpha} + \xi$$

$$p = \boldsymbol{x}' \boldsymbol{\gamma} - \frac{1}{\beta} \frac{dh}{dq} q + \eta$$

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

$$\overline{h} + \overline{\frac{dh}{dq}} (q - \overline{q}) \approx \beta p + \boldsymbol{x}' \boldsymbol{\alpha} + \xi$$

$$\overline{\frac{dh}{dq}} q \approx \beta p + \boldsymbol{x}' \boldsymbol{\alpha} + \xi - \overline{h} + \overline{\frac{dh}{dq}} \overline{q}$$

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

$$p \approx \boldsymbol{x}' \boldsymbol{\gamma} - \frac{1}{\beta} \frac{\tilde{dh}}{dq} q - \frac{1}{\beta} \left(\beta p + \boldsymbol{x}' \boldsymbol{\alpha} + \xi - \overline{h} + \overline{\frac{dh}{dq}} \overline{q} \right) + \eta$$

$$2p \approx \boldsymbol{x}' \boldsymbol{\gamma} + \frac{1}{\beta} \boldsymbol{x}' \boldsymbol{\alpha} - \frac{1}{\beta} \overline{h} + \frac{1}{\beta} \overline{\frac{dh}{dq}} \overline{q} - \frac{1}{\beta} \frac{\tilde{dh}}{dq} q + \eta + \frac{1}{\beta} \xi$$

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

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

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

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

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

$$h \approx \beta \left(\boldsymbol{x}' \boldsymbol{\gamma} - \frac{1}{\beta} \frac{\tilde{dh}}{dq} q - \frac{1}{\beta} \overline{\frac{dh}{dq}} q + \eta \right) + \boldsymbol{x}' \boldsymbol{\alpha} + \xi$$

$$h \approx \beta \boldsymbol{x}' \boldsymbol{\gamma} + \boldsymbol{x}' \boldsymbol{\alpha} - \frac{\tilde{dh}}{dq} q - \left(h - \overline{h} + \overline{\frac{dh}{dq}} \overline{q} \right) + \beta \eta + \xi$$

$$2h \approx \beta \boldsymbol{x}' \boldsymbol{\gamma} + \boldsymbol{x}' \boldsymbol{\alpha} - \frac{\tilde{dh}}{dq} q + \overline{h} - \overline{\frac{dh}{dq}} \overline{q} + \beta \eta + \xi.$$

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

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

Equations (C.6) and (C.7) provide an approximation for β.

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

$$\approx -\sqrt{\frac{\beta^2 Var\left(\eta + \frac{1}{\beta}\xi - \frac{1}{\beta}H^*\right)}{Var\left(\eta + \frac{1}{\beta}\xi - \frac{1}{\beta}H^*\right)}}$$

$$\approx \beta$$

QED.

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

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

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

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

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

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

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 (C.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 (C.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

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

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

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

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

C.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 (7), real solutions satisfy the condition:

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

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

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

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

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

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.

C.7 Proof of Proposition 5 (Non-Constant Marginal Costs)

Under the semi-linear marginal cost schedule of equation (13) and the assumption that $Cov(\xi, \eta) = 0$, the plim of the OLS estimator is equal to

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

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

$\xi + \left(\frac{1}{\beta}\frac{Cov(\xi,\lambda)}{Var(p^*)} - \frac{Cov(\xi,g(q))}{Var(p^*)}\right)p^*$. This implies

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

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

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

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

where

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

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

Therefore, we obtain the relations

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

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

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

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

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

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

$$= Cov(\xi^{OLS},g(q)) - \frac{1}{\beta}Cov(\xi^{OLS},\lambda).$$

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

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

and the following quadratic β.

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

QED.