A Comparison of Testing and Estimation of Firm Conduct

Lorenzo Magnolfi Christopher Sullivan*

January 2022

Abstract

Researchers both test and estimate structural models to learn firm conduct. As testing imposes candidate models suggested by economic theory, it is less demanding of the instruments. However, relative performance under misspecification depends on whether a candidate model approximates the truth.

JEL classification: L2, *Keywords:* firm conduct, testing, instruments, misspecification

1 Introduction

Key questions across fields depend on the nature of firm conduct (e.g., optimal taxation, monopsony power in labor markets, collusion). Learning conduct by regressing market outcomes (e.g., prices) on measures of market concentration (e.g., HHI) is problematic (see e.g., Berry, Gaynor, and Scott Morton, 2019). The alternative is to fully specify models of demand and firm costs, and either test or estimate models of firm conduct.

In markets for differentiated products, this latter approach assumes that the data for a set of products $j = 1, \dots, J$ in market t solve the stacked first order conditions to the firms' static profit maximization problems:[1]

$$\boldsymbol{p}_t = \boldsymbol{\Delta}_0(\boldsymbol{s}_t, \boldsymbol{p}_t) + \boldsymbol{c}_{0t},$$

where $\boldsymbol{\Delta}_0(\boldsymbol{s}_t, \boldsymbol{p}_t)$ is the true vector of markups, which depend on endogenous prices $\boldsymbol{p}_t$ and market shares $\boldsymbol{s}_t$, and $\boldsymbol{c}_{0t}$ is the vector of marginal costs.

Distinguishing the nature of conduct amounts to learning $\boldsymbol{\Delta}_0$ which the researcher can accomplish by relying on assumptions rooted in theory, and on the data. On a spectrum from least to most data-driven, researchers may assume a model of conduct, test a non-nested menu of possible models of conduct, or estimate parametric or nonparametric markup functions. Some research questions naturally lend themselves to one method. For instance, if a researcher wants to know whether firms compete in prices or quantities, testing is the appropriate tool. Instead, a researcher interested in measuring the evolution of markups would specify a flexible function $\boldsymbol{\Delta}(\boldsymbol{s}_t, \boldsymbol{p}_t)$ and, given sufficient variation, estimate it from data.

*Department of Economics, University of Wisconsin, Madison, WI 53706; Magnolfi (corresponding author): magnolfi@wisc.edu, Sullivan: cjsullivan@wisc.edu.

[1] Alternatively, the production function approach uses firms' cost minimization problems to infer markups. See De Loecker and Scott (2016) for a comparison.

However, settings where models of conduct are nested in a parametric specification of the markup function are amenable to both testing and estimation. Consider a common specification in the literature:

$$\boldsymbol{\Delta}(\theta) = \left[\Omega(\theta) \odot \left(\frac{\mathrm{d}\boldsymbol{s}_t}{\mathrm{d}\boldsymbol{p}_t}\right)'\right]^{-1} \boldsymbol{s}_t, \tag{1}$$

where we suppress the dependence of $\boldsymbol{\Delta}$ on $\boldsymbol{s}_t$ and $\boldsymbol{p}_t$. Different models m are described by values of $\theta = \theta_m$ which in turn define the internalization matrix Ω. Let Θ denote a continuum of models defining the support of θ. While Ciliberto and Williams (2014), Miller and Weinberg (2017), and Michel and Weiergraeber (2018) estimate $\theta_m \in \Theta$ to learn the model of collusion, Backus, Conlon, and Sinkinson (2021) and Duarte, Magnolfi, and Roncoroni (2020) instead test a discrete set of models $\mathcal{M} \subset \Theta$ to study common ownership and non-profit objectives, respectively.

We consider in this paper how a researcher should choose between estimation and testing in a setting where theory suggests two candidate models that are nested in a markup function. The two models correspond to two different values of a scalar θ.[2] The existing guidance in Nevo (1998) emphasizes an advantage to testing when the dimension of θ is large. In this case, estimation requires at least as many instruments as parameters, while testing a subset of models only requires one instrument. However, the applications referenced above impose additional structure on the internalization matrix, reflected in the scalar dimension of θ.

Using the results in Berry and Haile (2014) and Duarte, Magnolfi, Sølvsten, and Sullivan (2021) (DMSS), we compare the testing procedure developed in Rivers and Vuong (2002) (RV) to estimation from the perspective of weak instruments, and also under misspecification.[3] We extend the intuition in Nevo (1998) that testing is less demanding of the data to our setting by contrasting falsification to identification. Relative performance under misspecification depends on how the researcher specifies the set of models, $\mathcal{M}$. If economic theory is informative in forming the menu of models, testing can conclude for a model closer to the truth than estimation, potentially even the true one. Our results can help applied researchers to evaluate the merits of each method.

2 Inference with Weak Instruments

Testing and estimation rely on the same econometric framework. We assume that firms' costs in market t are $\boldsymbol{c}_{0t} = \mathbf{w}_t\gamma + \boldsymbol{\omega}_{0t}$, where $\mathbf{w}_t$ and $\boldsymbol{\omega}_{0t}$ are uncorrelated vectors of observed

[2] Our formal claims are easy to extend to higher dimensional parameters, and more than two candidate models.

[3] While alternative procedures have been used to test conduct, DMSS show they have poor inferential properties under misspecification, so we focus solely on RV. See the Online Appendix for the formulation of the RV test.

and unobserved cost shifters respectively.[4] It is useful to eliminate $\mathbf{w}$ from the model, akin to keeping the observable part of marginal cost constant. For any variable $\boldsymbol{y}$, we define the residualized variable $y = \boldsymbol{y} - \mathbf{w}E[\mathbf{w}'\mathbf{w}]^{-1}E[\mathbf{w}'\boldsymbol{y}]$.

Testing and estimation further require instruments z_t that are excluded from (1) and uncorrelated with ω_{0t}.[5] The moment condition $E[z_{jt}\omega_{0jt}] = 0$ thus characterizes the true model. For all $\theta \in \Theta$, let $\omega_t(\theta) = p_t - \Delta(\theta)$, and use the moment $E[z_{jt}\omega_{jt}(\theta)]$ to construct the GMM lack-of-fit $Q(\theta)$. Instruments z falsify a model m if $E[z_{jt}\omega_{jt}(\theta_m)] \neq 0$; global identification of θ requires $Q(\theta) > 0$ for all $\theta \neq \theta_0$, the parameter value corresponding to the true model.

Estimation fails when identification breaks down. Testing, performed with RV, fails in the population when neither of the two models can be falsified by the instruments, causing the RV test to be degenerate (Duarte et al., 2021). To inform the choice between these two methods, we formalize the relationship between degeneracy and lack of identification:

Claim 1. *Suppose that the RV test for testing models $m = 1, 2$ is degenerate. Then, global identification fails. Conversely, even if θ_0 is not identified, the RV test for models $m = 1, 2$ may not be degenerate.*

The degeneracy of RV testing is thus a special case of lack of identification, complementing the result in Nevo (1998). While Nevo (1998) argues that testing is less onerous than estimation as it requires fewer exclusion restrictions, in our setting θ is scalar. However, the intuition that testing is less demanding of the instruments still holds. As estimation requires the instruments to differentiate a continuum of models, stronger instruments are needed than for testing two models.

For ease of exposition we state our result in terms of fixed models in the population. In finite samples, weak instruments can cause a near failure of identification or falsifiability, leading to inferential problems. Thus, it is paramount to diagnose weak instruments. Although informal tests have been proposed (Michel and Weiergraeber, 2018), no statistic yet exists to diagnose weak instruments for estimation in non-linear GMM (Stock and Wright, 2000). Instead, DMSS propose an effective F-statistic and provide critical values which can be used to evaluate the quality of the inference from RV testing.

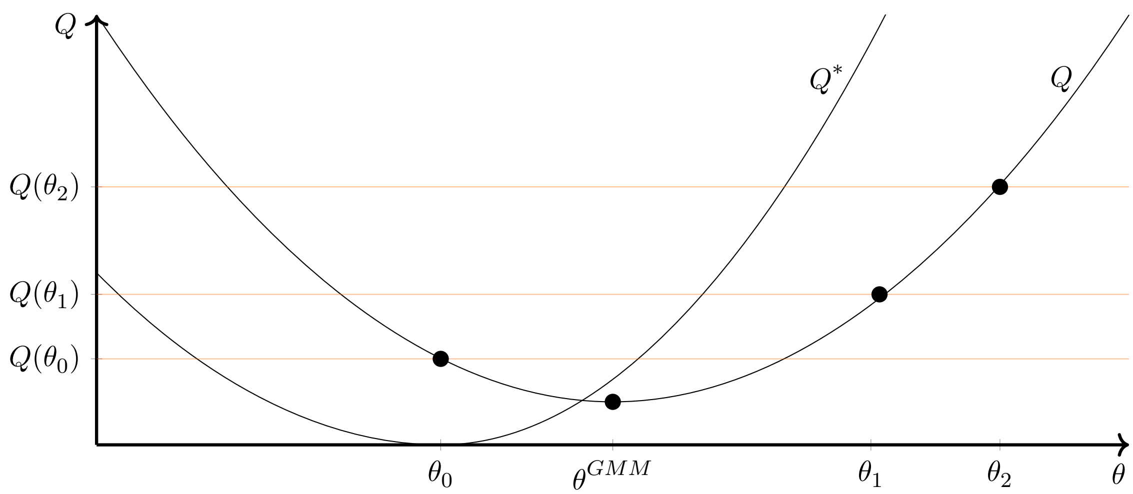
[4]This assumes marginal costs are constant in quantity and a linear index in $\mathbf{w}$. Appendix B of DMSS shows these assumptions are not necessary for our arguments.

[5]We maintain in this note that the researcher has precommitted to a set of instruments z_t formed with appropriate sources of variation (see e.g. Berry and Haile, 2014). DMSS discuss how testing can be extended to use all available sources of exogenous variation without requiring the researcher to precommit.

3 Misspecification

With strong instruments and no misspecification, estimation is consistent, while testing concludes in favor of the true model if it is included in $\mathcal{M}$. However, without nonparametric modeling of demand, cost, and markups functions, misspecification is likely. We illustrate the consequences of testing and estimating conduct with misspecification in Figure 1. Q^* represents the population GMM objective function without misspecification, defined over Θ and minimized at $\theta = \theta_0$. We maintain misspecification is consequential so that the GMM objective function used for both estimation and testing is $Q \neq Q^*$, which is minimized at the pseudo-true value $\theta^{GMM} \neq \theta_0$.

FIGURE 1: Estimation and Testing of Conduct



We illustrate the effect of misspecification on estimation and RV testing. Q^* and Q denote the GMM objective function without and with misspecification respectively.

A researcher can estimate conduct, inferring θ^{GMM} from a large sample. With misspecification, estimation is inconsistent and there is no clear interpretation for θ^{GMM}. Alternatively, a researcher could specify two models of conduct to test. The RV test always concludes in favor of the model for which predicted markups (markups projected on instruments) are closer to the true predicted markups (see DMSS). Thus, the quality of the results depends on the menu of models considered by the researcher. If $\mathcal{M} = \{\theta_0, \theta_1\}$, $Q(\theta_0) < Q(\theta_1)$, and RV asymptotically concludes for the true model. However, if $\mathcal{M} = \{\theta_1, \theta_2\}$, RV concludes for θ_1, which is farther from the truth than θ^{GMM}. We summarize this discussion in the following claim:

Claim 2. *Suppose that misspecification is consequential. Then, estimation of conduct is inconsistent, while RV testing can conclude in favor of the true model. However, the test may also conclude in favor of a model that is farther from the truth than the estimated one.*

In finite sample, additional inferential considerations support testing over estimation. Although progress is being made in locally misspecified settings (e.g., Hansen and Lee, 2021), inference with standard two-step GMM methods under global misspecification may be misleading (Hall and Inoue, 2003). Thus, the confidence intervals obtained from GMM estimation in globally misspecified models should not be used to evaluate conduct. Instead, the inferential properties of RV are well understood under misspecification (see DMSS).

4 Discussion

Testing conduct uses external information by forming a menu of models. This has two consequences. First, testing places a smaller burden on the data, and the RV test may not be degenerate even if θ_0 is not identified. Second, the implications for distinguishing conduct under misspecification are nuanced. If the economic theory used to form the set of candidate models is informative, testing can conclude for a model closer to the truth than estimation, potentially even the true one.

5 Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Backus, M., C. Conlon, and M. Sinkinson (2021): "Common Ownership and Competition in the Ready-To-Eat Cereal Industry," Working paper.

Berry, S., M. Gaynor, and F. Scott Morton (2019): "Do increasing markups matter? lessons from empirical industrial organization," *Journal of Economic Perspectives*, 33, 44–68.

Berry, S. and P. Haile (2014): "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 82, 1749–1797.

Ciliberto, F. and J. Williams (2014): "Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry," *RAND Journal of Economics*, 45, 764–791.

De Loecker, J. and P. Scott (2016): "Estimating Market Power: Evidence from the US Brewing Industry," NBER Working Paper # 22957.

Duarte, M., L. Magnolfi, and C. Roncoroni (2020): "The Competitive Conduct of Consumer Cooperatives," Working paper.

Duarte, M., L. Magnolfi, M. Sølvsten, and C. Sullivan (2021): "Testing Firm Conduct," Working paper.

Hall, A. and A. Inoue (2003): "The Large Sample Behaviour of the Generalized Method of Moments Estimator in Misspecified Models," *Journal of Econometrics*, 114, 361–394.

Hansen, B. and S. Lee (2021): "Inference for iterated GMM under misspecification," *Econometrica*, 89, 1419–1447.

Michel, C. and S. Weiergraeber (2018): "Estimating Industry Conduct in Differentiated Products Markets," Working paper.

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

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

Rivers, D. and Q. Vuong (2002): "Model Selection Tests for Nonlinear Dynamic Models," *Econometrics Journal*, 5, 1–39.

Stock, J. and J. Wright (2000): "GMM with weak identification," *Econometrica*, 68, 1055–1096.

Online Appendix

Here we expand on the testing environment in the article and introduce some notational conventions. Letting i index a product j in market t, the GMM objective functions introduced in the article are defined as $Q(\theta) = g(\theta)'Wg(\theta)$ where $g(\theta) = E[z_i(p_i - \Delta_i(\theta))]$ and $W = E[z_i z_i']^{-1}$ is a weight matrix. We define $\hat{Q}(\theta)$ as the sample analog of $Q(\theta)$. An important object for testing and estimation are predicted markups, defined as $\Delta^z_m = z\Gamma_m$ where the projection coefficient is $\Gamma_m = E[z'z]^{-1}E[z'\Delta(\theta_m)]$. The GMM objective function can then be expressed in terms of the MSE in predicted markups for a model, or $Q(\theta_m) = E[z_i(p_i - \Delta_i(\theta_m))]'E[z_i z_i']^{-1}E[z_i(p_i - \Delta_{mi})] = E[(\Delta^z_{0i} - \Delta^z_{mi})^2]$. If $\mathcal{M} = \{\theta_1, \theta_2\}$, the RV test statistic is $T^{\text{RV}} = \sqrt{n}(\hat{Q}(\theta_1) - \hat{Q}(\theta_2))/\hat{\sigma}_{\text{RV}}$, where σ^2_{RV} is an estimator of the asymptotic variance of the numerator of the test statistic. The GMM estimator is defined as $\hat{\theta}^{GMM} = \arg\min_\theta \hat{Q}(\theta)$.

We also maintain throughout the following assumptions:

Assumption 1. $\boldsymbol{z}_i$ is a vector of d_z excluded instruments, so that $E[\boldsymbol{z}_i\omega_{0i}] = 0$.

Assumption 1 maintains that the instruments are exogenous. We also maintain that the standard assumptions for GMM estimation hold, see e.g., assumptions A.1-A.6 in Hall and Inoue (2003).

Proof of Claim 1. By Corollary 1 in DMSS, RV is degenerate if and only if $E[z_i(p_i - \Delta_i(\theta_m)] = 0$ for $m = 1, 2$ so that neither model is testable. In turn, if two values of θ satisfy the moment condition, $Q(\theta_1) = Q(\theta_2) = 0$ and identification fails.

To prove the converse, a counterexample suffices. Suppose that identification fails because there exists a range of parameter values $\bar{\Theta} \subset \Theta$ such that $Q(\theta) = 0$ for all $\theta \in \bar{\Theta}$. As long as either $\theta_1 \notin \bar{\Theta}$ or $\theta_2 \notin \bar{\Theta}$, then RV is not degenerate. □

To prove Claim 2 we maintain that misspecification is consequential in the following sense:

Assumption 2. Let $\theta^{GMM} = plim\hat{\theta}^{GMM}$ be such that $\theta^{GMM} \neq \theta_0$.

For convenience, we also maintain that the functional form of markups nests the true model of conduct in the absence of misspecification of demand or cost. Let $\Delta^*(\theta_0)$ indicate the markups for the true model in the absence of misspecification and $\Delta(\theta)$ the markup function imposed by the researcher. If $\Delta(\theta)$ is misspecified, then $\Delta^*(\theta_0) \neq \Delta(\theta_0)$. Furthermore, Δ^{*z}_0 denotes the predicted markups for the true model without misspecification.

Proof of Claim 2. For the purposes of this proof, we maintain that θ is identified. By Claim 1, this implies that RV is not degenerate for any two models over the support of θ.

Consider the following two cases illustrated in Figure 1:

Case *(i)*: suppose that $\mathcal{M} = \{\theta_0, \theta_1\}$. Then, by Lemma 2 in DMSS, with probability approaching one as $n \to \infty$, RV rejects the null of equal fit in favor of the true model as $E[(\Delta^{*z}_{0i} - \Delta^z_{0i})^2] < E[(\Delta^{*z}_{0i} - \Delta^z_{1i})^2]$. This holds as under misspecification $E[(\Delta^{*z}_{0i} - \Delta^z_{mi})^2] = E[z_i(p_i - \Delta_i(\theta_m))]' E[z_i z_i']^{-1} E[z_i(p_i - \Delta_i(\theta_m))] = Q(\theta_m)$.

Case *(ii)*: suppose instead that the researcher tests models $\mathcal{M} = \{\theta_1, \theta_2\}$ corresponding to $\theta_1, \theta_2 \neq \theta_0$. Then, without loss of generality, if $E[(\Delta^{*z}_{0i} - \Delta^z_{1i})^2] < E[(\Delta^{*z}_{0i} - \Delta^z_{2i})^2]$ as in Figure 1, the RV test will conclude asymptotically in favor of model $m = 1$. However, $E[(\Delta^{*z}_{0i} - \Delta^z_{1i})^2] > E[(\Delta^{*z}_{0i} - \Delta^z_{GMMi})^2] > 0$ so that the expected squared distance between the predicted markups for model 1 and the true predicted markups is larger than the expected squared distance between the predicted markups Δ^z_{GMMi} implied by θ^{GMM} and the true predicted markups, even assuming Assumption 2 holds.

□

Appendix References

Hall, A. and A. Inoue (2003): "The Large Sample Behaviour of the Generalized Method of Moments Estimator in Misspecified Models," *Journal of Econometrics*, 114, 361–394.