Production Function Identification Under Imperfect Competition*

Daniel A Ackerberg — UT Austin

Jan De Loecker — KU Leuven and CEPR

June 23, 2024

Abstract

The presence of imperfect competition introduces distinct challenges when estimating production functions. We start by highlighting that some existing approaches to production function estimation cannot completely abstract away from the presence of imperfect competition in the product market. We then extend these existing approaches to accommodate some additional oligopoly models commonly used in empirical work by using a sufficient statistic approach, and show that the presence of such strategic interactions has important benefits in that they introduces additional exogenous variation that can help identify production functions. We study how to optimally leverage this exogenous variation, both with and without direct data on a firm's competitors, and use Monte-Carlo experiments to 1) verify that the existence of strategic interactions can identify production functions that would not otherwise be identified, and 2) assess the extent to which what applied researchers observe about competition affects the precision of estimates based on this variation.

Keywords: *Production Functions; Imperfect Competition; Sufficient Statistic; Identification*

*Contact the authors at `jan.deloecker@kuleuven.be` and `daniel.ackerberg@utexas.edu`. Thanks to Allan Collard-Wexler, Mitsuru Igami, Qingsong Pan, and seminar participants at KU Leuven, UBC, and UC-Berkeley for comments. De Loecker acknowledges support from FWO's Odysseus Program.

1 Imperfect competition: challenges and benefits

The estimation of production functions is an important step in a host of economic analyses ranging from identifying drivers of productivity growth to the study of market power. There exists a variety of approaches, e.g. the "proxy variable" literature originating with Olley and Pakes (1996, OP) and Levinsohn and Petrin (2003, LP) which we refer to throughout as "OP/LP-style approaches", and the dynamic panel approach of Blundell and Bond (2000, BB). While in principle these approaches only require fully specifying the *technological* relationship between output and inputs, they do put restrictions, some more explicitly than others, on the underlying operating environment of producers. This paper studies these restrictions in the context of imperfect competition and potential strategic interactions in the product market. While much of the literature has been carried out without explicitly taking a stand on the underlying market structure in the product market, we highlight that these approaches to production function estimation cannot (or should not) completely abstract away from the presence of imperfect competition.[1]

After providing a brief summary of the key aspects of these papers, we summarize the extent to which existing applications of these techniques, in particular those based on OP/LP-style approaches, are consistent with imperfect competition. We then show how they can be extended to feasibly allow more general models of imperfect competition by deriving *sufficient statistic* results that can be used to control for the presence of strategic interactions among producers. The models covered include homogeneous good quantity setting competition (e.g. Cournot) and Bertrand-Nash price setting under a (nested) logit model of demand - these are models of oligopoly commonly used in applied work. We show that in these cases, OP/LP-style approaches can use a modified "first stage" inversion equation based on a low dimensional sufficient statistic. A fundamental contribution of the OP/LP based literature is that their semiparametric approaches allow production functions to be estimated without fully specifying other aspects of firms maximization problems. Our extensions to imperfect competition for this set of models inherit this propery, i.e. we are able to estimate production functions without fully specifying models of demand, input choice, and in some cases the nature of competition. We then argue that if one wants to consider more general models of strategic interactions, one can alternatively make additional technological restrictions, e.g. a restricted linear process for productivity as in the BB dynamic panel (DP) approach, or an approach discussed in Ackerberg, Caves, and Frazer (2015, ACF) and De Loecker and Scott (2022) based on a Leontief assumption that we call a

[1]We note that other approaches may be able to though reliance on additional assumptions. In particular, as detailed by De Loecker and Syverson (2021), the factor share approach only requires observing cost shares, but imposes assumptions of constant returns to scale and that all inputs in production are variable factors.

technology control (TC) approach.

While the above illustrates how imperfect competition is non-trivial for the first stage of OP/LP-style approaches, it on the other hand can provide additional exogenous variation that can help in the second stage of these approaches (as well as the TC and DP approaches), More specifically, it introduces a new set of valid instruments - the productivity shocks and fixed factors of a firm's competitors - that shift the residual demand curve facing a firm and thus can generate exogenous variation in that firms' use of variable inputs. This can be thought of as a supply side analogue to the common use of competitive instruments in demand estimation, e.g. Bresnahan (1987) and Berry, Levinsohn, and Pakes (1995). We show how the existence of these "oligopoly instruments" can in fact break the important non-identification result of Gandhi, Navarro, and Rivers (2020, GNR), and, applying Chamberlain (1985), show that there are interesting considerations in considering optimal use of these instruments. Moreover, we show that lagged variable inputs can be valid instruments in the case of strategic interactions among producers, similarly breaking the non-identification result of GNR under perfect competition. We also note an important implication of this, i.e. that utilizing imperfect competition as an source of exogenous variation does not necessarily require defining markets or even observing firms' sets of competitors. These issues are relevant not only for OP/LP-style approaches, but also the TC and DP approaches.

We conclude with a Monte-Carlo study and some extensions of our results. The Monte-Carlo study verifies that oligopoly instruments can identify production functions (that would otherwise not be identified as per GNR), and assesses the precision of estimates obtained through different competitive instruments, in particularly comparing the case where competitors are observed versus when they are not observed. In the extensions, we assess a key limitation of competitive instruments in that they cannot generally provide independent exogenous variation for multiple variable inputs, assess the applicability of our results to situations in which only firm revenues (and not physical quantities) can be reliably observed, and consider the possibility of imperfect competition also being present in input markets.

Lastly, we want to emphasize that our focus is on situations where one's primary goal is to estimate production functions and one *does not have the exogenous price variation to additionally estimate a demand system*. In other words, we envision a "standard" production dataset primarily composed of data on inputs and outputs. If one additionally has the exogenous price variation and other needed data to sufficient fully specify and estimate a demand system, more could likely relax some of the substantial assumptions we make here, e.g. regarding unobserved demand shocks or product characteristics and models of demand and competition. While recent years have seen an increase in the availability of datasets that have the variation necessary to estimate both production and demand, they still unfortunately constitute a small minority.

2 Setup

Consider a firm j producing output (Q_{jt}) in a market/period t (which we use to index both different geographic markets and the same geographic market across time) using observed variable and fixed inputs V_{jt} (e.g. labor) and F_{jt} (e.g. capital), with an unobserved (to the econometrician) productivity shock ω_{jt}. We consider a production function of the form

$$Q_{jt} = Q(V_{jt}, F_{jt}; \theta) \exp(\omega_{jt}) \exp(\epsilon_{jt}) \tag{1}$$

or in logs (represented by lower case and used interchangeably through the paper as per convenient)

$$q_{jt} = q(v_{jt}, f_{jt}; \theta) + \omega_{jt} + \epsilon_{jt} \tag{2}$$

where ϵ_{jt} is classical measurement error in q_{jt} that is mean independent of all other variables in the model (though possibly correlated over time). f_{jt} can represent a vector of multiple fixed inputs, though for now we assume that the variable input v_{jt} is scalar. Important issues regarding multivariate v_{jt} are discussed at the end of the paper.

There are two aspects of OP/LP-style and dynamic panel approaches that we review here because they are particularly relevant for extending the approaches to imperfect competition in product markets (one can consult the surveys of Ackerberg, Berry, Benkard, and Pakes (2007, ABBP) and De Loecker and Syverson (2021, DLS) for more extensive discussion). First is the semi-parametric "first stage" inversion that is a crucial step for OP/LP-style approaches, but not dynamic panel approaches. Second are "second stage" moment conditions based on timing and information set assumptions used by OP/LP-style approaches - as discussed in ACF and elsewhere these are very similar to the moment conditions used in the dynamic panel approach of BB. We now discuss each of these aspects in turn, as well as a technology control approach that also relies on the second stage moment conditions.

2.1 OP/LP-Style First and Second Stages

While we describe them generically as OP/LP-style approaches, we actually restrict attention to LP's proposed use of V_{jt} (or equivalently, v_{jt}) as the "first stage" control variable (the original OP approach instead uses investment). This restriction is crucial because the sufficient statistic results here do not apply when investment is used as the control variable (fortunately a large proportion of the literature does in fact use V_{jt} as the control, i.e. where our results apply). This highlights one of the advantages pointed out by LP regarding using V_{jt} as the control variable - optimal choice of V_{jt} can often be characterized by considering a relatively simple static first order condition, while investment is inherently dynamic and more complicated.

More specifically, LP derive conditions under which firms' optimal choice of variable input V_{jt} can be written as

$$V_{jt} = V_t(\omega_{jt}, F_{jt}) \tag{3}$$

where V_t is strictly monotonic in the scalar unobservable ω_{jt}. One example of these conditions is the standard assumption of price taking in factor markets, and perfectly competitive output markets, though as we discuss below, LP also consider more general conditions. The importance of the scalar unobservable (plus strict monontonicity) assumption is that it implies that any two firms in the same market t with the same observed V_{jt} and F_{jt} must have the same ω_{jt}. Note that since V is indexed by t, (3) is consistent with firms facing different output or input prices in different markets, and these do not need to be observed by the econometrician.

Given conditions under which (3) holds, the first stage proceeds by inverting for ω_{jt} and substituting into (2), i.e.

$$q_{jt} = q(v_{jt}, f_{jt}; \theta) + V_t^{-1}(V_{jt}, F_{jt}) + \epsilon_{jt} \tag{4}$$

Since this equation only has the single econometric unobservable ϵ_{jt} that by assumption is mean independent of V_{jt} and F_{jt} (and their logged versions v_{jt} and f_{jt}), there are no endogeneity problems in (4). However, OP/LP-style approaches treat the inverse function V_t^{-1} non-parametrically, and hence $q(v_{jt}, f_{jt}; \theta)$ is not identified from solely this equation. But one can non-parametrically regress q_{jt} on v_{jt} and f_{jt} and this first stage, while not identifying $q(v_{jt}, f_{jt}; \theta)$, does produce estimates of the sum $q(v_{jt}, f_{jt}; \theta) + V_t^{-1}(V_{jt}, F_{jt})$ and the residuals ϵ_{jt}.[2]

[2] More commonly the non-parametric component of (4) is written in log form $q_{jt} = q(v_{jt}, f_{jt}; \theta) + v_t^{-1}(v_{jt}, f_{jt}) + \epsilon_{jt}$, but this is equivalent given V_t^{-1} (v_t^{-1}) is treated non-parametrically.

Denoting the estimates of the residuals as $\widehat{\epsilon}_{jt}$, we next consider the "second stage" of OP/LP-style approaches. For clarity we present this under the assumptions that the production function is Cobb-Douglas and that ω_{jt} follows the AR(1) process $\omega_{jt} = \rho\omega_{jt} + \xi_{jt}$ over time, though the results can be generalized. Given this, we can write the production function as

$$\widetilde{q}_{jt} \equiv q_{jt} - \widehat{\epsilon}_{jt} = \theta_0 + \theta_f f_{jt} + \theta_v v_{jt} + \omega_{jt} \tag{5}$$

with the single unobservable ω_{jt}[3].

Second stage estimation then proceeds under the assumption that the innovation in the productivity shock process satisfies

$$E\left[\xi_{jt} \mid I_{jt-1}\right] = 0 \tag{6}$$

where I_{jt-1} is firm j's information set at time $t-1$. Various different timing and information set assumptions have been made in the literature, but a typical one is that $f_{jt} \in I_{jt-1}$, but that $v_{jt} \notin I_{jt-1}$ (though $v_{jt} \in I_{jt}$). This corresponds to the assumption that firms' do not observe the innovation ξ_{jt} until time t, plus a standard interpretation of variable and fixed inputs - e.g. that variable inputs can be chosen up to the time of production (t), while fixed inputs are predetermined, i.e. need to be chosen prior to production time (e.g. at $t-1$).

Using (5), the moment condition (6) can be rewritten in terms of observables as[4]

$$E\left[\left(\widetilde{q}_{jt} - \theta_0 - \theta_f f_{jt} - \theta_v v_{jt}\right) - \rho\left(\widetilde{q}_{jt-1} - \theta_0 - \theta_f f_{jt-1} - \theta_v v_{jt-1}\right) \mid I_{jt-1}\right] = 0 \tag{7}$$

Using moment conditions like these to estimate the production parameters (and ρ) is the basis of the second stage of most OP/LP-style approaches, though as discussed below, there are some cases where these moment conditions have been shown to not identify the parameters (GNR).

[3]Pakes and Olley (1995) and Hahn, Liao, and Ridder (2018) discuss how the fact that $\widehat{\epsilon}_{jt}$ is estimated in a prior stage affects inference, an issue we abstract away from here.

[4]While to conserve notation we use a single subscript t to describe markets that are different either geographically or temporally, when t and $t-1$ appear in the same equation they should be taken to indicate the same geographic market in sequential time periods.

2.2 Dynamic Panel Approach

As alluded to above, the dynamic panel (DP) approach to estimating production functions (BB) utilizes similar moment conditions, though since these approaches do not utilize a first stage to estimate the ϵ_{jt}, for this model they instead would use the moment condition

$$
\begin{aligned}
& E\left[\xi_{jt} + \epsilon_{jt} - \rho\epsilon_{jt-1} \mid I_{jt-1}\right] && (8)\\
= \; & E\left[(q_{jt} - \theta_0 - \theta_f f_{jt} - \theta_v v_{jt}) - \rho\left(q_{jt-1} - \theta_0 - \theta_f f_{jt-1} - \theta_v v_{jt-1}\right) \mid I_{jt-1}\right] = 0 && (9)
\end{aligned}
$$

where again, the exact composition of I_{jt-1} depends on the specific timing and information set assumptions one makes. As discussed in, e.g. ACF, Shenoy (2021), DLS, Ackerberg (2023), and Doraszelski and Jaumandreu (2023), an important advantage of the DP approach is that by avoiding the OP/LP first stage, it does not require the scalar unobservable and strict monotonicity assumptions of that first stage. Doraszelski and Jaumandreu (2023) advocate using the DP approach under imperfect competition for exactly this reason. On the other hand, unlike OP/LP-style approaches, standard DP approaches do not generalize to more flexible 1st order productivity shock processes.[5]

2.3 Technology Control Approach

Another approach that avoids the OP/LP first stage inversion (also at the cost of additional assumptions) is mentioned in ACF and De Loecker and Scott (2022). It relies on the assumption that the production function is Leontief in a second variable input M_{jt} and of the form

$$Q_{jt} = \min\left\{Q(V_{jt}, F_{jt}; \theta)\Omega_{jt}, G(M_{jt})\right\}\Xi_{jt} \tag{10}$$

M_{jt} for exanple, could be a material input that is utilized proportionally to output. In this case, assuming there are no unutilized inputs (i.e. the firm sets $Q(V_{jt}, F_{jt}; \theta)\Omega_{jt} = G(M_{jt})$), one can non-parametrically regress q_{jt} on M_{jt} (or $m_{jt} = \ln(M_{jt})$) to obtain estimates of the measurement errors, $\widehat{\epsilon}_{jt}$. With these in hand, one can directly form the moments (7). We denote this last approach a *technology control* (TC) approach as

[5] As detailed in the aforementined papers, the tradeoff between OP and DP approaches can loosely be described as 1) DP approaches do not require the scalar unobservable and strict monotonicity assumptions necessary to identify $\widehat{\epsilon}_{jt}$ in the OP first stage, versus, 2) DP approaches require stronger assumptions on first order productivity shock process, i.e. while OP style approaches can straightforwardly be generalized to $\omega_{jt} = g(\omega_{jt}) + \xi_{jt}$, it is harder to generalize DP approaches from a simple AR(1) ω_{jt} process (though with additional differencing DP approaches can add an additional fixed effect α_i to the product function).

it is based primarily on an additional assumption (Leontief) regarding production technology. In relation to OP/LP-style approaches, the tradeoff is that by making this additional Leontief assumption on M_{jt}, the TC approach, like the DP approach, does not require the scalar unobservable and strict monotonicity assumptions required to invert (3).[6]

2.4 Discussion

Given the above setup, we can layout how this paper proceeds regarding assessing the impact of imperfect competition on these three different estimation approaches - OP/LP, DP, and TC. The paper is divided into two parts. The first part corresponds to the first stage inversion of the variable input demand function $V_{jt} = V_t(\omega_{jt}, F_{jt})$. Being mindful of the scalar unobservable assumption necessary for this inversion, we consider the extent to which this can be applied (or extended) to allow imperfect competition. As described above, this is *only relevant for OP/LP-style approaches* (since the DP and TC approaches do not require the inversion).

In contrast, the second part of the paper concerns how imperfect competition affects the second stage moment conditions, e.g. $E\left[\xi_{jt} \mid I_{jt-1}\right] = 0$ or $E\left[\xi_{jt} + \epsilon_{jt} - \rho\epsilon_{jt-1} \mid I_{jt-1}\right]$, and is relevant for *all three approaches*. In short we show how imperfect competition and strategic interactions can be helpful in identifying production function parameters using these moments. In some cases it enables one to avoid the non-identification result of GNR. We then study different aspects of various versions of the competitive instruments we advocate, including their efficiency properties, and illustrate some of their characteristics with Monte-Carlo experiments.

3 First Stage Inversion with Imperfect Competition

3.1 Canonical OP/LP Proxy Variable

It has previously been recognized (e.g. OP, LP, ABBP, DLW, and ACF) that the basic factor demand function described in the prior section

$$V_{jt} = V_t(\omega_{jt}, F_{jt}) \tag{11}$$

[6]The TC approach might also be particularly sensitive to measurement error in M (though of course all the discussed estimators are likely inconsistent in this context, see, e.g. Collard-Wexler and De Loecker (2022)). Of course, if one has the choice of multiple material inputs, they can choose the one they think least prone to measurement error.

is not consistent with all forms of imperfect competition. This equation characterizes firms' optimal choice of the variable input, and in general, with imperfect competition, firm j's optimal choice of V_{jt} will not only depend on their own ω_{jt} and F_{jt}, but also productivity shocks and fixed input levels of other firms in market j.

However, as noted by the aforementioned work, equation (11) can hold to the extent that competition is *symmetric* enough (up to observables). As one example, suppose that firms j in market t are monopolistic competitors, e.g. who each produce and sell in their own geographic submarkets with identical downward sloping demand curves (and face identical input prices). Since there is no strategic interaction between firms in market t, and since each firm in market t is facing the same input and output market conditions (note that these conditions can vary across markets t since V is indexed by t), equation (11) will hold. The assumption of symmetry across these monopolistic competitors can be relaxed to the extent that any asymmetries are observed. For example, suppose that in market t, all firms j face the same demand curve up to some vector Z_{jt} that is observed by the econometrician. This can be accommodated by simply adding Z_{jt} into the function, i.e.

$$V_{jt} = V_t(\omega_{jt}, F_{jt}, Z_{jt}) \tag{12}$$

Since Z_{jt} is assumed observed, the scalar unobservable assumption still holds. By the same logic, Z_{jt} can also be used to capture any differences in input markets faced by firms j in market t, again to the extent to which they are observed (e.g. each firm faces a different price of the variable input V_{jt} and these different prices are observed). Under these assumptions, any two firms at t who have the same ω_{jt} and F_{jt} and face the same output and input market conditions (measured by Z_{jt}) will optimally choose the same V_{jt}.

This monopolistic competition environment highlights the importance of a key contribution of OP regarding the power of their semiparametric approach. Specifically, the fact that OP/LP treat V_t nonparametrically means that the production function coefficients can be estimated *without fully specifying* the form of the demand curves faced by the monopolistic competitors, nor estimating those demand curves. It only requires that these demand curves be the same for all j (at a given t) up to observables Z_{jt}. It is a distinct advantage to be able to estimate some structural parameters (in this case production function parameters) without needing to commit to a parametric model of other aspects of the environment.[7]

Equation (11) (or (12)) can also hold in models with actual strategic interactions between firms in a market t, as long as there is sufficient symmetry. Suppose that all the firms in market t compete in a

[7]Similarly, when an investment control variable is used (as in the original OP), the semiparametric approach allows one to estimate production function parameters without needing to fully specify the dynamic investment problem of firms.

homogeneous good Cournot game, choosing V_{jt} (or equivalently Q_{jt}) conditional on fixed levels of (ω_{jt}, F_{jt}) for all j. Note that since the firms in market t may have different (ω_{jt}, F_{jt}), this is a Cournot game where firms have heterogeneous variable cost curves (the cost curve for j is determined by the production function evaluated at (ω_{jt}, F_{jt})). Because firms are symmetric up to (ω_{jt}, F_{jt}) in this model, equation (11) will generally hold[8]. The key here is that in the standard OP/LP approach, V_t is typically indexed by market t. Therefore what (11) requires is that in a given market t, any two firms j and j' with the same ω_{jt} and F_{jt} optimally choose the same level of the variable input. In this case this is true because firms are symmetric on the demand side of a homogeneous good Cournot model. Hence the set of competitors faced by j (which includes j') is identical to the set of competitors faced by j' (which includes j). In other words, in this model the t subscript on $V_t(.)$ is sufficient to capture the effect of strategic interactions on factor demand. This was first pointed out by OP.[9] The same logic can hold for simple differentiated product Bertrand-Nash models that are sufficiently symmetric, e.g. consider (12) where Z_{jt} contains all the product characteristics of firm j's product.

There is, however, a conceptual problem with the above approach of indexing the function $V_t(.)$ by t to implicitly allowing for models with strategic interaction like Cournot and Nash-Bertrand. Recall that OP/LP-style approaches treat the function $V_t(.)$ (i.e. $V_t^{-1}(.)$) non-parametrically in order to avoid specifying other parts of the model (e.g. the demand curve, dynamic optimization). Hence, when one is indexing V by market t, econometric consistency requires the number of observations (firms) in *each market* t to go to infinity. But a large number of firms in a given market is not typically what one has in mind when thinking about Cournot and differentiated product Nash-Bertrand models of imperfect competition. More practically, for markets we typically think of as being imperfectly competitive - perhaps with less than 10-20 firms in each market - it is going to be very hard to credibly estimate a separate $V_t(\omega_{jt}, F_{jt})$ function non-parametrically for each market t.

As pointed out by DLS, instead of estimating a separate $V_t(\omega_{jt}, F_{jt})$ for each market t, an alternative is to remove the t index but put enough arguments into $V(.)$ such that it holds *across* markets (again, across both time and space). In particular, they consider a function of the form

$$V_{jt} = V(\omega_{jt}, F_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, Z_t) \tag{13}$$

where $\boldsymbol{\omega}_{-jt}$ and $\boldsymbol{F}_{-jt}$ are vectors of the productivity shocks for all the competitors to firm j in market t

[8] Under certain regularity conditions, e.g. that reaction functions have unique maximizers.

[9] See pages 1273 and 1277.

(boldface indicates vectors/matrices of these variables for all firms $k \neq j$). Assuming any differences in demand and supply conditions across markets t are captured by observables Z_t, this will generally hold in homogenous good Cournot equilibria across markets[10]. Intuitively, (13) could be hypothetically derived by solving out the system of FOCs for V_{jt} in the Cournot game. It will depend on all primitives including the demand curve. But as usual, with the OP/LP-style approach to production function estimation, one does not need to solve out the system of equations since (the inverse of) V is treated non-parametrically in the first stage[11].

Similarly, in a differentiated product Nash-Bertrand game, one can consider the function

$$V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt}, Z_t) \tag{14}$$

where X_{jt} contains observed product characteristics for firm j's product (and $\boldsymbol{X}_{-jt}$ contains product characteristics of all j's competitors). If all other determinants of demand are symmetric across firms and markets t (e.g. logit errors), then (14) will generally hold in Nash-Bertrand equilibria *across* markets t. Again, this equation could hypothetically be derived from the system of first order conditions, but the OP/LP semi-parametric approach allows one to avoid doing this.

On the other hand, using (13) or (14) is likely to be challenging in practice, and perhaps as a result, we are not aware of a paper that actually applies this approach. This is due to at least two reasons. First, there are J (the number of firms in market t) unobservables in (13) (or (14)) - hence one would need to consider a J dimensional inversion:

$$\begin{gathered}
\omega_{1t} = V^{-1}(V_{1t}, F_{jt}, \mathbf{V}_{-1t}, \boldsymbol{F}_{-1t}, Z_t) \\
. \\
. \\
\omega_{Jt} = V^{-1}(V_{Jt}, F_{jt}, \mathbf{V}_{-Jt}, \boldsymbol{F}_{-Jt}, Z_t)
\end{gathered}$$

i.e. invert firm j's productivity shock as a function of not only firm j's choice of variable input, but also all other firms' choices of variable input (V_{-j}) (and conditional on their respective fixed inputs (F_{-j})). Second, and perhaps more importantly, even if this multivariate inversion exists, the number of arguments of the

[10] Again under regularity conditions, e.g. either that the primitives are such that equilibrium is always unique, or that if there are multiple equilibria, the same equilibria is played across all markets t.

[11] This is particularly beneficial in this case because general Cournot games with heterogeneous, non-constant marginal cost curves typically do not have analytic solutions.

non-parametric function V^{-1} will tend to be large (at least twice the number of firms per market), and thus this may be hard to operationalize in practice when there are more than a few firms in each market.

3.2 Discussion and Our Sufficient Statistic Approach

In sum, we have argued that while existing OP/LP-style approaches can in theory accomodate (at least some forms of) imperfect competition in two ways - 1) by indexing the function V_t by t, or 2) by including the entire "state vector" of market participants in V - both may have problems in practice. Our goal in the first half of this paper is to develop approaches that avoid these problems. In particular, we look for *reduced dimension* analogues of (13) (or (14), i.e. models of imperfect competition, e.g. Cournot, differentiated product Nash-Bertrand, where we can derive functions of the form:

$$V_{jt} = V(\omega_{jt}, F_{jt}, Z_{jt}, Z_t) \text{ (or } V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Z_{jt}, Z_t)) \tag{15}$$

i.e. where 1) the V's are not indexed by t, i.e. they can be treated non-parametrically using data *across all markets*, and 2) Z_{jt} is of reduced dimension, i.e. does not have dimension proportional to the number of competitors. In other words, the function does not need to contain vectors of all other firms' productivities and fixed factors ($\boldsymbol{\omega}_{-jt}$ and $\boldsymbol{F}_{-jt}$, plus characteristics $\boldsymbol{X}_{-jt}$ in differentiated product Bertrand).

In the next three subsections, we derive such functions in three sets of models. In particular, we show that in homogenous good quantity setting models and Nash-Bertrand Logit/Nested Logit models, the total output of firm j's competitors, $Q_{-jt} = \sum_{k \neq j} Q_{kt}$, is a *sufficient statistic* to describe firm j's input demand.[12] In other words, in equilibrium, if firm j's competitors are producing Q_{-jt}, it doesn't matter how many competitors firm j is facing, or the distribution of productivities of those competitors - that firm must always be choosing the same V_{jt}. This means we can express

$$V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, Z_t), \tag{16}$$

across markets and reduce the dimension of the non-parametric object being modeled.[13] We note that

[12] This is somewhat reminiscent of sufficient statistics as used in the theoretical games literature (see, e.g. Nocke and Schutz (2018) and cites therein), but for quite different purposes - in our case the goal is to obtain a sufficient statistic for an inversion equation that holds in equilibria across different markets in a dataset.

[13] For simplicity of notation we assume that Q_{-jt} is the only element of Z_{jt}. However what follows trivially extends to allowing other elements of Z to vary across j (as long as they are observed). In other words, instead of $V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, Z_t)$ we could have $V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, Z_{jt})$. This allows, for example, different firms j in a given market t to face different input prices (again, as long as they are observed by the econometrician).

Doraszelski and Jaumandreu (2023, DJ) have independently derived results related to some of our sufficient statistic results.[14]

3.3 Sufficient statistic in a Homogeneous Good Quantity Setting Model

We first show our sufficient statistic result in a homogenous goods Cournot model, and then generalize to more general (homogenous good) quantity setting models. Suppose inverse demand in each market t is given by

$$P\left(Q_t, Z_t\right),$$

where Q_t is the total quantity produced and Z_t captures demand shifters across markets. Note that because our goal is to use the same $V\left(\right)$ across markets t, Z_t will need to be observed. Otherwise the scalar unobservable assumption crucial to the OP/LP-style approaches would be violated. However, we do not need to fully specify the inverse demand function $P\left(\right)$. This is because $P\left(\right)$ enters the econometric model through its implied input demand function $V\left(\right)$, which is treated non-parametrically.

Supposing each firm j's production function is given by

$$Q_{jt} = Q(V_{jt}, F_{jt}; \theta) \exp(\omega_{jt}), \tag{17}$$

where Q is strictly increasing in its arguments, we can think about a Cournot game conditional on $(\omega_{jt}, F_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt})$.[15] In other words, consider a set of firms with *fixed* values of fixed inputs and productivity shocks choosing levels of a single variable input V_{jt} in a Nash Equilibrium. While traditional Cournot games are posed with quantity Q_{jt} as the choice variable, under mild conditions the production function (17) is a one-to-one mapping between V_{jt} and Q_{jt}, so considering V_{jt} as the choice variable is equivalent.

Given $(\omega_{jt}, F_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, Z_t)$ and choices $\mathbf{V}_{-j,t}$ of its competitors, firm j's variable profit function is

$$\Pi_j = P\left(Q(V_{jt}, F_{jt}; \theta) \exp(\omega_{jt}) + \sum_{k \neq j} Q(V_{kt}, F_{kt}; \theta) \exp(\omega_{kt}), Z_t\right) Q(V_{jt}, F_{jt}; \theta) \exp(\omega_{jt}) - P_V(Z_t) V_{jt}$$

[14] See their Section 6. DJ also derive expressions for the bias resulting from the econometrician omitting a relevant argument of V (e.g. Z_{jt} or Z_t) - i.e. violating the scalar unobservable assumption - we assume throughout the paper that this is not the case.

[15] Note the absence of the econometrician measurement errors ϵ_{jt} in the production function. This is fine for the purposes of the theoretical derivations in this and the next two subsections. When we return to empirical application, we reintroduce the measurement errors, and show how they are quite important.

where we also allow the price of the competitively purchased input to depend on observed market specfic variables in Z_t.[16] Defining the scalar valued

$$Q_{-j}(V_{-jt}, \omega_{-jt}, F_{-jt}) = \sum_{k \neq j} Q(V_{kt}, F_{kt}; \theta) \exp(\omega_{jt}).$$

the Nash F.O.C. for firm j equates marginal revenue to marginal cost, i.e.

$$P\left(Q(V_{jt}, .) + Q_{-j}(V_{-j,t}, .), Z_t\right) + P'\left(Q(V_{jt}, .) + Q_{-j}(V_{-j,t}, .), Z_t\right) Q(V_{jt}, .) = \frac{P_V(Z_t)}{\frac{\partial Q(V_{jt}, .)}{\partial V_{jt}}} \tag{18}$$

where for compactness we define $Q(V_{jt}, .) = Q(V_{jt}, F_{jt}; \theta) \exp(\omega_{jt})$. Firm j's optimal choice of V_{jt} conditional on $(V_{-jt}, \omega_{jt}, F_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, Z_t)$ satisfies this F.O.C. We make the following high level assumption

Assumption 1 *The primitives of the model $P()$, $Q()$, and $P_V()$ are such that for all $(V_{-jt}, \omega_{jt}, F_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, Z_t)$ there is a unique $0 < V_{jt} < \infty$ that solves (18)*

This assumption reflects relatively standard conditions, e.g. quasiconcavity of the profit function, that ensure a well-defined reaction function (rather than correspondence). Given this, we can show the following Theorem holds regarding the equilibria of this Cournot game across different markets:

Theorem 1 *Under Assumption (1), consider any two markets t and t' - the first with environmental variables $(\omega_{jt}, F_{jt}, \omega_{-jt}, F_{-jt}, Z_t)$ and the second with environmental variables $(\omega_{jt'}, F_{jt'}, \omega_{-jt'}, F_{-jt'}, Z_{t'})$. Suppose that $\omega_{jt} = \omega_{jt'}, F_{jt} = F_{jt'}$, and $Z_t = Z_{t'}$, but that (ω_{-jt}, F_{-jt}) and $(\omega_{-jt'}, F_{-jt'})$ are not equal in the two markets (the vectors ω_{-j} and F_{-j} may even have a different number of elements, i.e. firm j may have a different number of competitors in each market). Consider an equilibrium in each of these markets, and suppose that, in these respective equilibria, the variable inputs chosen by firms other than j, i.e. V_{-jt} and $V_{-jt'}$, generate equilibrium total quantities produced by firms other than j that are equal (i.e. $Q_{-j}(V_{-jt}, \omega_{-jt}, F_{-jt}) = Q_{-j}(V_{-jt'}, \omega_{-jt'}, F_{-jt'})$). Then it must be the case that in these two respective equilibria, $V_{jt} = V_{jt'}$*

Proof. Consider the FOC (1) at the two respective equilibria. By supposition, $\omega_{jt} = \omega_{jt'}, F_{jt} = F_{jt'}$, and $Z_t = Z_{t'}$. (ω_{-jt}, F_{-jt}) is not equal to $(\omega_{-jt'}, F_{-jt'})$, but these variables only enter the FOCs through the

[16] Again, $P_V(Z_t)$ need not be specified since it will enter the production function through the non-parametric V^{-1}, and can be generalized to depend on observed Z_{jt}'s that are firm specific.

scalar valued function $Q_{-j}(\cdot)$, and by supposition $Q_{-j}(V_{-jt}, \omega_{-jt}, F_{-jt}) = Q_{-j}(V_{-jt'}, \omega_{-jt'}, F_{-jt'})$. Hence, $V_{jt} \neq V_{jt'}$ contradicts Assumption (1) that there is a unique V_{jt} for any $(V_{-jt}, \omega_{jt}, F_{jt}, \omega_{-jt}, F_{-jt}, Z_t)$. ■

Loosely speaking, Theorem (1) says that from firm j's perspective, the structure of its competition $-j$ only "matters" through the total quantity produced by its competitors, i.e. Q_{-j}. In other words, firm j will do the same thing regardless of whether it facing just a few, efficient, competitors vs. many, more inefficient, competitors - as long as in equilibrium the total production of those respective sets of competitors is the same.

Most importantly for our purposes, Theorem (1) implies that if we look across different markets t with the same Z_t, firms operating in those markets with the same ω_{jt}, F_{jt} - and whose competitors are producing total quantity $Q_{-jt} = \sum_{k \neq j} Q_{kt}$ - must in equilibrium be choosing the same level of V_{jt}. In other words, there is a function V such that in equilibria across these markets t, the following relationship holds:

$$V_{jt} = V(\omega_{jt}, F_{jt}, Q_{-jt}, Z_t), \tag{19}$$

where the *scalar* Q_{-jt} is a sufficient statistic for describing the equilibrium relationship between j's competitors actions and j's action. Note that it is not quite appropriate to describe this as an input demand function - because Q_{-jt} is an equilibrium object. However, (19) is an relationship that holds in equilibrium across markets t, and thus we will be able to treat it non-parametrically in the first stage of an OP/LP-style approach. And relative to (13), it is of reduced dimension and thus more manageable empirically.

Lastly, note that while the above results were derived for the Cournot-Nash equilibria of a homogenous product quantity-setting game, they generalize to a more general homogenous product quantity setting game with a "conduct parameter" Θ. In such games, the first order condition is

$$P\left(Q(V_{jt}, .) + Q_{-j}(V_{-jt}, .), Z_t\right) + \Theta(Q(V_{jt}, .), Q_{-j}(V_{-jt}, .), Z_t) P'\left(Q(V_{jt}, .) + Q_{-j}(V_{-jt}, .), Z_t\right) Q(V_{jt}, .) = \frac{P_V(Z_t)}{\frac{\partial Q(V_{jt}, .)}{\partial V_{jt}}} \tag{20}$$

so the arguments above go through. Note that this allows the conduct parameter to depend on $Q(V_{jt}, .), Q_{-j}(V_{-jt}, .)$, and Z_t, and that the market level observables Z_t can contain the number of firms in the market N_t. As well known (see, e.g. Bresnahan (1982), Berry and Haile (2014), Backus, Conlon, and Sinkinson (2021), and Duarte, Magnolfi, Solvesten, and Sullivan (2024)), this indexes a set of models including both Cournot Nash and Perfect Collusion, and as long as Assumption (1) holds, the relationship (19) continues to hold across markets in equilibrium.[17] Note that in this case, not only does the semiparametric aspect of OP/LP-style approaches allow one to estimate a production function without fully specifying the demand function $P(Q_t, Z_t)$, but also without specifying (or estimating) the conduct parameter function $\Theta(Q(V_{jt}, .), Q_{-j}(V_{-j,t}, .), Z_t)$. (19) holds regardless of this function, as long as the relevant variables are observed.

3.4 Sufficient statistic in the Logit Nash-Bertrand Model

The above result is probably not that surprising given we know in the homogeneous good Cournot model firms' reaction functions only depend on their competitors actions through the scalar Q_{-jt}. But we now show that Q_{-jt} can also serve as a sufficient statistic in simple Nash Bertrand games with logit based product differentiation. Consider the following consumer level "logit" utility function

$$\begin{aligned} U_{ijt} &= f(X_{jt}, P_{jt}, Z_t) + \varepsilon_{ijt} \\ U_{i0t} &= \varepsilon_{i0t} \end{aligned}$$

where P_{jt} is the price charged by firm j in market t, and X_{jt} is a vector of characteristics of good j in market t. ε_{ijt} and ε_{i0t} are standard logit errors, and Z_t again includes observed market level variables that may affect demand. With a continuum of consumers i we obtain the standard logit market share formula for firm j

$$S_{jt} = \frac{e^{f(X_{jt}, P_{jt}, Z_t)}}{1 + \sum_{k \in t} e^{f(X_{kt}, P_{kt}, Z_t)}}$$

and if market size is given by $M(Z_t)$, the quantity demanded for firm j is:

$$Q_{jt}(P_{jt}, \boldsymbol{P}_{-jt}, X_{jt}, \boldsymbol{X}_{-jt}, Z_t) = M(Z_t) \frac{e^{f(X_{jt}, P_{jt}, Z_t)}}{1 + \sum_{k \in t} e^{f(X_{kt}, P_{kt}, Z_t)}} \tag{21}$$

Note that this assumes markets are identical up to the observables $(\boldsymbol{X}_t, \boldsymbol{P}_t, Z_t)$ ($\boldsymbol{X}_t$ and $\boldsymbol{P}_t$ contain characteristics and prices of all products in market t) in this model. So, for example, there cannot be unobserved product characteristics. It is also a pure logit model, i.e. there are no random coefficients. On the other hand, because we do not need to fully specify the function f (again, this is because of the semi-parametric approach of the OP/LP first stage), residual demand in this model can be a very flexible function of P_{jt}.

[17] Lastly, note that while we describe this as a homogenous product quantity setting game, the fact that we allow Z_t to enter the inverse demand curve actually allows products to potentially be heterogeneous across markets (though not across firms within a market) to the extent that this is can be measured by Z_t. One could potentially allow the production function to depend on elements of Z_t as well.

As in the Cournot model, we consider a production function of the form

$$Q_{jt} = Q(V_{jt}, F_{jt}; \theta)\exp(\omega_{jt}) \tag{22}$$

that is strictly increasing in its arguments.[18] With these primitives we consider equilibria of the following Nash-Bertrand game. Conditional on F_{jt}, ω_{jt}, X_{jt}, and Z_t for all firms j in market t, we assume that firms simultaneously choose prices P_{jt}. Given any vector of prices, quantities are given by (21) $\forall j$, and each firm j must then purchase enough of the variable input V_{jt} to produce that quantity. This level is given by the inverse of the production function

$$V_{jt} = Q^{-1}(Q_{jt}, \omega_{jt}, F_{jt}; \theta) \tag{23}$$

which exists assuming the production function is strictly increasing in V_{jt}.

Given $(\omega_{jt}, F_{jt}, X_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt}, Z_t)$ and choices $\boldsymbol{P}_{-jt}$ of its competitors, firm j's (variable) profit function is

$$\begin{aligned}
\Pi_{jt} &= P_{jt}Q_{jt}\left(P_{jt}, \boldsymbol{P}_{-jt}, X_{jt}, \boldsymbol{X}_{-jt}, Z_t\right) - P_V\left(Z_t\right)Q^{-1}(Q_{jt}\left(P_{jt}, \boldsymbol{P}_{-jt}, X_{jt}, \boldsymbol{X}_{-jt}, Z_t\right), \omega_{jt}, F_{jt}; \theta) \quad (24)\\
&= P_{jt}M(Z_t)\frac{e^{f(X_{jt},P_{jt},Z_t)}}{1+\sum_{k\in t} e^{f(X_{kt},P_{kt},Z_t)}} - P_V\left(Z_t\right)Q^{-1}(M(Z_t)\frac{e^{f(X_{jt},P_{jt},Z_t)}}{1+\sum_{k\in t} e^{f(X_{kt},P_{kt},Z_t)}}, \omega_{jt}, F_{jt})\\
&= P_{jt}M(Z_t)\frac{e^{f(X_{jt},P_{jt},Z_t)}}{1+e^{f(X_{jt},P_{jt},Z_t)}+\Phi_{-jt}} - P_V\left(Z_t\right)Q^{-1}(M(Z_t)\frac{e^{f(X_{jt},P_{jt},Z_t)}}{1+e^{f(X_{jt},P_{jt},Z_t)}+\Phi_{-jt}}, \omega_{jt}, F_{jt})
\end{aligned}$$

where

$$\Phi_{-jt} = \sum_{k\in t/j} e^{f(X_{kt},P_{kt},Z_t)}$$

This illustrates that firm j's (variable) profit function only depends on $(\boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt})$ through the scalar Φ_{-jt}, which can be interpreted as a measure of the "strength" of competition firm j is facing in equilibrium. For example, Φ_{-jt} could be large because firm j is facing a few competitors with high X_{kt} and low P_{kt}, or it could be large because firm j is facing many competitors (though with "worse" X_{kt}'s and P_{kt}'s).

Given this structure of the first order condition we can show that Φ_{-jt} can serve as a sufficient statistic in this Bertrand model, . Again we start with a high level assumption ensuring that the primitives (e.g.

$f\left(\right), Q\left(\right)$, and $P_V\left(\right)$) are such that there is a unique maximum of this profit maximization problem:

[18]Note that F_{jt} can include X_{jt}, which implicitly allows the production function to depend on observed product characteristics (see, e.g. Hahn (2023)).

Assumption 2 *The primitives of the model* $f()$, $Q()$, *and* $P_V()$ *are such that for all* $(P_{-jt}, \omega_{jt}, F_{jt}, X_{jt}, \omega_{-jt}, F_{-jt}, X_{-jt}, Z_t)$ *there is a unique* $0 < P_{jt} < \infty$ *that maximizes (24)*

Then the following Theorem holds in equilibria of this Bertrand-Nash game across markets:

Theorem 2 *Under Assumption (2), consider any two markets* t *and* t' *- the first with environmental variables* $(\omega_{jt}, F_{jt}, X_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt}, Z_t)$ *and the second with environmental variables* $(\omega_{jt'}, F_{jt'}, X_{jt'}, \boldsymbol{\omega}_{-jt'}, \boldsymbol{F}_{-jt'}, \boldsymbol{X}_{-jt'}, Z_{t'})$. *Suppose that* $\omega_{jt} = \omega_{jt'}, F_{jt} = F_{jt'}, X_{jt} = X_{jt'}$, *and* $Z_t = Z_{t'}$, *but that* $(\boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt})$ *and* $(\boldsymbol{\omega}_{-jt'}, \boldsymbol{F}_{-jt'}, \boldsymbol{X}_{-jt'})$ *are not equal in the two markets (again, firm* j *may have a different number of competitors in each market so these vectors are of different dimension). Consider an equilibrium in each of these markets, and suppose that, in these respective equilibria, the prices chosen by firms other than* j, *i.e.* $\boldsymbol{P}_{-j,t}$ *and* $\boldsymbol{P}_{-j,t'}$, *are such that* $\Phi_{-jt} = \Phi_{-jt'}$. *Then it must be the case that in these two respective equilibria,* $P_{jt} = P_{jt'}, Q_{jt} = Q_{jt'}$, *and* $V_{jt} = V_{jt'}$.

Proof. Consider the profit function at the two respective equilibria. By supposition, $\omega_{jt} = \omega_{jt'}, F_{jt} = F_{jt'}, X_{jt} = X_{jt'}$, and $Z_t = Z_{t'}$. $(\boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt})$ is not equal to $(\boldsymbol{\omega}_{-jt'}, \boldsymbol{F}_{-jt'}, \boldsymbol{X}_{-jt'})$, but these variables only enter profits through the scalar valued function Φ, and by supposition $\Phi_{-jt} = \Phi_{-jt'}$. Hence, $P_{jt} \neq P_{jt'}$ would contradict Assumption (2). Since $P_{jt} = P_{jt'}$, it must additionally be the case that $Q_{jt} = Q_{jt'}$ (since in the logit model residual demand is strictly decreasing in price), and therefore that $V_{jt} = V_{jt'}$ given the production function is assumed strictly increasing in V. ■

Again loosely speaking, Theorem (2) shows that from firm j's perspective, the structure of its competition $-j$ only "matters" through the scalar sufficient statistic Φ_{-jt}. It also implies that across different markets t with the same Z_t, firms operating in those markets with the same $\omega_{jt}, F_{jt}, X_{jt}$ - and whose competitors $(\boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt})$ and P_{-jt} are such that in equilibrium Φ_{-jt} is the same - must 1) in equilibrium be choosing the same P_{jt}, and therefore 2) also using the same amount of the variable input V_{jt}. In other words, there is a function $V()$ such that in equilibria across these markets, the following relationship holds:

$$V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, \Phi_{-jt}, Z_t) \tag{25}$$

i.e. the scalar Φ_{-jt} is a sufficient statistic for describing the equilibrium relationship between firm j's choice of variable input and its competitors' actions. The Theorem also implies that there is a function P such that the following relationship holds

$$P_{jt} = P(\omega_{jt}, F_{jt}, X_{jt}, \Phi_{-jt}, Z_t) \tag{26}$$

Since P_{jt} is the strategic variable that is directly chosen in this price-setting game, (26) can be interpreted as a sufficient statistic representation of a reaction function - i.e. firm j's optimal price as a function of its own characteristics, market characteristics, and the scalar Φ_{-jt} summarizing the impact of other firms prices (and X's) on j's profits. This reaction function will be useful momentarily.

While Φ_{-jt} is a natural sufficient statistic in the logit model for the level of competition faced by firm j, (25) is unfortunately not that useful from the perspective of OP/LP-style approaches since they rely on a scalar unobservable assumption. This is a problem for (25) since Φ_{-jt} will typically not be observable, since 1) it depends on the unknown function $f()$, and 2) it depends on prices P_{-jt}, which are typically not be observed in a canonical production function dataset on inputs and outputs. Hence, we next consider conditions under which Q_{-jt} can serve as an alternative sufficient statistic - that is, where there is an alternative function $V()$ such that

$$V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, Z_t) \tag{27}$$

To show this, first note that if $y = f(a, b)$ and $c = g(a, b)$, then if g is strictly monotonic in b there exists a function $\overline{f}$ such that $y = \overline{f}(a, c)$ $(\equiv f(a, g^{-1}(a, c)))$. In other words, if y is a function of a and b, and there is a monotonic relationship between b and c (for any a), then y can also be expressed as a function of a and c. In our situation, the standard logit formulas imply

$$\begin{aligned} Q_{-jt} &= M(Z_t)\frac{\sum_{k\in t/j} e^{f(X_{kt},P_{kt},Z_t)}}{1+\sum_{k\in t} e^{f(X_{kt},P_{kt},Z_t)}} \\ &= M(Z_t)\frac{\Phi_{-jt}}{1+e^{f\left(X_{jt},P_{jt},Z_t\right)}+\Phi_{-jt}} \end{aligned}$$

Moreover, based on the implication of Theorem (2), we know that in equilibrium, across markets j, the following relationship holds

$$Q_{-jt} = M(Z_t)\frac{\Phi_{-jt}}{1+e^{f(X_{jt},P(\omega_{jt},F_{jt},X_{jt},\Phi_{-jt},Z_t),Z_t)}+\Phi_{-jt}} \tag{28}$$

where $P(\omega_{jt}, F_{jt}, X_{jt}, \Phi_{-jt}, Z_t)$ is the sufficient statistic reaction function (26). (28) implies that in equilibrium, Q_{-jt} can be expressed as a function of $(\omega_{jt}, F_{jt}, X_{jt}, Z_t)$ and Φ_{-jt}. Hence, the following assumption

Assumption 3 *For all* $(\omega_{jt}, F_{jt}, X_{jt}, Z_t)$ *and* Φ_{-jt}

$$\frac{\partial \left(\frac{\Phi_{-jt}}{1+e^{f\left(X_{jt}, P(\omega_{jt}, F_{jt}, X_{jt}, \Phi_{-jt}, Z_t), Z_t\right)}+\Phi_{-jt}} \right)}{\partial \Phi_{-jt}} > 0$$

suffices for $V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, Z_t)$ to hold, i.e. allowing us to replace Φ_{-jt} with Q_{-jt} as our scalar sufficient statistic to capture the relationship between j's competition and j's choice of variable input.

Assumption (3) is a relatively high level assumption. Whether it holds will depend on the primitives, i.e. the utility function $f(X_{jt}, P_{jt}, Z_t)$, the production function $Q(V_{jt}, F_{jt})$, and the input price function $P_V(Z_t)$. Ideally one would construct more primitive conditions under which this would hold, but this is complicated by the fact that our goal is to estimate production functions without fully specifying demand (it depends on the unspecified $f()$), and moreover, even for a specific $f()$, note that reaction functions do not generally have a closed form in this model.

However, examining (28) suggests that Assumption (3) is intuitive and may not be all that stringent. Recall that $\Phi_{-jt} = \sum_{k\in t/j} e^{f(X_{kt}, P_{kt}, Z_t)}$ is a scalar index summarizing the utility provided by firms other than j. Note that if $P(\omega_{jt}, F_{jt}, X_{jt}, \Phi_{-jt}, Z_t)$ hypothetically did not depend on Φ_{-jt}, then Assumption (3) would certainly hold - i.e. if the "utility" provided by firms other than j increases (i.e. Φ_{-jt} increases), then holding P_j constant, Q_{-jt} must increase. Of course, in a Nash-Bertrand logit model, prices tend to be strategic complements. So when Φ_{-jt} increases, we would expect $P(\omega_{jt}, F_{jt}, X_{jt}, \Phi_{-jt}, Z_t)$ to decrease, and this could potentially change the sign of the derivative. Essentially, what Assumption (3) requires is that this reaction function response is not too large. If the $-j$ "utility" index increases (e.g. due to lower prices or better X for firms other than j) - which will naturally increase Q_{-jt} - the downward price response of j cannot be large enough to *reverse* that Q_{-jt} increase. This is the assumption required for our sufficient statistic result.

3.5 Sufficient statistic in the Nested Logit Nash-Bertrand Model

Lastly, we derive a sufficient statistic result for the Nested Logit Nash-Bertrand Model. We do this in the context of a single level nested logit model, in which we obtain a two-vector sufficient statistic, but it is generalizeable to more levels (in which case the sufficient statistic is equal to the number of levels plus one).

The utility functions are just a nested logit version of our logit model, i.e.

$$\begin{aligned} U_{ijt} &= f\left(X_{jt}, P_{jt}, Z_t\right) + \varsigma_{ig_jt} + (1 - \sigma\left(Z_t\right))\varepsilon_{ijt} \\ U_{i0t} &= \varsigma_{i0t} + (1 - \sigma\left(Z_t\right))\varepsilon_{i0t} \end{aligned}$$

where nests are indexed by $g \in G_t$ and g_{jt} represents the nest that firm j is in in market t. We assume the outside alternative is in its own group $g = 0$, and the nesting parameter $\sigma\left(Z_t\right)$ can vary across markets depending on observables. As above, $f\left(X_{jt}, P_{jt}, Z_t\right)$ is unspecified and thus this model can capture fairly general shapes of demand.

Define

$$\begin{aligned} D_{gt} &= \sum_{k \in J_{gt}} e^{f(X_{kt}, P_{kt}, Z_t)/(1-\sigma(Z_t))}, \\ \Phi_{g(-j)t} &= D_{g_jt} - e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma(Z_t))}, \text{ and} \\ \Phi_{G(-g_j)t} &= \sum_{(g \neq g_{jt}) \in G_t} D_{gt}^{(1-\sigma(Z_t))} \end{aligned}$$

where J_{gt} represents the set of products in nest g in market t. $\Phi_{g(-j)t}$ is a scalar measure of the attractiveness of the other products *in* j's nest, and $\Phi_{G(-g_j)t}$ is a scalar measure of the attractiveness of all the products *outside* of j's nest. In this model, residual demand for firm j in nest g_j is

$$\begin{aligned} & Q_{jt}\left(P_{jt}, P_{-jt}, X_{jt}, X_{-jt}, Z_t\right) \\ = \; & M(Z_t) \frac{e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma(Z_t))}}{D_{g_jt}^{\sigma(Z_t)}(1 + \sum_{g \in G_t} D_{gt}^{(1-\sigma(Z_t))})} \\ = \; & M(Z_t) \frac{e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma(Z_t))}}{\left(e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma)} + \Phi_{g(-j)t}\right)^{\sigma(Z_t)} \left(1 + \left(e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma(Z_t))} + \Phi_{g(-j)t}\right)^{1-\sigma(Z_t)} + \Phi_{G(-g_j)t}\right)} \end{aligned}$$

Thus, analagously to the logit model, under

Assumption 4 *The primitives of the model* $f\,()$, $Q\,()$, $P_V\,()$, *and* $\sigma\,()$ *are such that for all* $(P_{-jt}, \omega_{jt}, F_{jt}, X_{jt}, \omega_{-jt}, F_{-jt}, X_{-jt}, Z_t)$ *there is a unique* $0 < P_{jt} < \infty$ *that maximizes profits.*

we have the following Theorem:

Theorem 3 *Under Assumption (4), consider any two markets t and t' - the first with environmental variables $(\omega_{jt}, F_{jt}, X_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt}, Z_t)$ and the second with environmental variables $(\omega_{jt'}, F_{jt'}, X_{jt'}, \boldsymbol{\omega}_{-jt'}, \boldsymbol{F}_{-jt'}, \boldsymbol{X}_{-jt'}, Z_{t'})$. Suppose that $\omega_{jt} = \omega_{jt'}, F_{jt} = F_{jt'}, X_{jt} = X_{jt'}$, and $Z_t = Z_{t'}$, but that $(\boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt})$ and $(\boldsymbol{\omega}_{-jt'}, \boldsymbol{F}_{-jt'}, \boldsymbol{X}_{-jt'})$ are not equal in the two markets (again, firm j may have a different number of competitors in each market so these vectors are of different dimension). Consider an equilibrium in each of these markets, and suppose that, in these respective equilibria, the prices chosen by firms other than j, i.e. $\boldsymbol{P}_{-j,t}$ and $\boldsymbol{P}_{-j,t'}$, are such that $\Phi_{g(-j)t} = \Phi_{g(-j)t'}$ and $\Phi_{G(-g_j)t} = \Phi_{G(-g_j)t'}$. Then it must be the case that in these two respective equilibria, $P_{jt} = P_{jt'}, Q_{jt} = Q_{jt'}$, and $V_{jt} = V_{jt'}$.*

Proof. Considering the profit function at the two respective equilibria, proof is analagous to proof of (2) ■

The Theorem implies that the following relationships hold across all markets

$$\begin{aligned} P_{jt} &= P(X_{jt}, \omega_{jt}, F_{jt}, \Phi_{g(-j)t}, \Phi_{G(-g_j)t}, Z_t), \qquad \text{and} \\ V_{jt} &= V(X_{jt}, \omega_{jt}, F_{jt}, \Phi_{g(-j)t}, \Phi_{G(-g_j)t}, Z_t) \end{aligned} \tag{29}$$

i.e. we now have two scalars, $\Phi_{g(-j)t}$ and $\Phi_{G(-g_j)t}$ that are sufficient statistics for P_{jt} and V_{jt}. Again, this is not sufficient for an OP/LP-style inversion since $\Phi_{g(-j)t}$ and $\Phi_{G(-g_j)t}$ are not observed, so we need additional assumptions to obtain a usable function to invert. Standard nested logit formulas also imply

$$\begin{aligned} & Q_{g(-j)t}\left(P_{jt}, P_{-jt}, X_{jt}, X_{-jt}, Z_t\right) \\ = \; & M(Z_t) \frac{\Phi_{g(-j)t}}{\left(e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma)} + \Phi_{g(-j)t}\right)^{\sigma} \left(1 + \left(e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma)} + \Phi_{g(-j)t}\right)^{1-\sigma} + \Phi_{G(-g_j)t}\right)} \end{aligned}$$

and

$$\begin{aligned} Q_{G(-g_j)t}\left(P_{jt}, P_{-jt}, X_{jt}, X_{-jt}, Z_t\right) &= M(Z_t) \frac{\sum_{(g \neq g_{jt}) \in G_t} D_{gt}^{(1-\sigma)}}{1 + \sum_{g \in G_t} D_{gt}^{(1-\sigma)}} \\ &= M(Z_t) \frac{\Phi_{G(-g_j)t}}{1 + \left(e^{f(X_{jt}, P_{jt}, Z_t)/(1-\sigma)} + \Phi_{g(-j)t}\right)^{1-\sigma} + \Phi_{G(-g_j)t}} \end{aligned}$$

and combining these with (29) implies that

$$Q_{g(-j)t} = M(Z_t)\frac{\Phi_{g(-j)t}}{\left[\begin{array}{c}\left(e^{f\left(X_{jt},P(X_{jt},\omega_{jt},F_{jt},\Phi_{g(-j)t},\Phi_{G(-g_j)t},Z_t),Z_t\right)/(1-\sigma)}+\Phi_{g(-j)t}\right)^{\sigma-1} \\ \cdot\left(1+\left(e^{f\left(X_{jt},P(X_{jt},\omega_{jt},F_{jt},\Phi_{g(-j)t},\Phi_{G(-g_j)t},Z_t),Z_t\right)/(1-\sigma)}+\Phi_{g(-j)t}\right)^{1-\sigma}+\Phi_{G(-g_j)t}\right)\end{array}\right]} \tag{30}$$

and

$$Q_{G(-g_j)t} = M(Z_t)\frac{\Phi_{G(-g_j)t}}{1+\left(e^{f\left(X_{jt},P(X_{jt},\omega_{jt},F_{jt},\Phi_{g(-j)t},\Phi_{G(-g_j)t},Z_t),Z_t\right)/(1-\sigma)}+\Phi_{g(-j)t}\right)^{1-\sigma}+\Phi_{G(-g_j)t}} \tag{31}$$

hold across equilibria.

Given primitives (and the implied equilibrium $P()$ function), (30) and (31) define a mapping from $\Re^2 \rightarrow \Re^2$

$$M_{NL}: \left(\Phi_{g(-j)t}, \Phi_{G(-g_j)t}\right) \rightarrow \left(Q_{g(-j)t}, Q_{G(-g_j)t}\right)$$

conditional on $(X_{jt}, \omega_{jt}, F_{jt}, Z_t)$. If the primitives are such that M_{NL} is invertible, then , then the following holds across equilibria

$$V_{jt} = V(X_{jt}, \omega_{jt}, F_{jt}, Q_{g(-j)t}, Q_{G(-g_j)t}, Z_t)$$

which we can use as the basis for a OP/LP-style first stage inversion. As in the logit model, a reason why the mapping would not be invertible would be "extreme" reactions, i.e. if in a market where, in equilibrium, attractiveness of competitors is higher (either $\Phi_{g(-j)t}$ or $\Phi_{G(-g_j)t}$), j's price negative response is strong enough to lower $Q_{g(-j)t}$ or $Q_{G(-g_j)t}$.

3.6 OP/LP Style Inversion with Sufficient Statistics

In the above three sets of models, we have derived functions that capture the effects of competitors through limited dimension summary statistics. However, there is an additional issue that must be addressed in order to invert and utilize these functions in an OP/LP-style first stage. A presumption of this literature is that measured output is contaminated by measurement error ϵ_{jt}. If this is the case for firm j, then it is presumably also the case for firm j's competitors. Thus, the sufficient statistic Q_{-jt} (or $Q_{g(-j)t}$ and $Q_{G(-g_j)t}$) will also be measured with error.

To address this issue, we now explicitly distinguish between true output Q^*_{jt} (and Q^*_{-jt}) and measured output Q_{jt} (and Q_{-jt}). We do this in the context of the Quantity Setting and Nash-Bertrand Logit models to conserve notation, but it generalizes to the Nested Logit model. Recalling the (log) production function

$$q_{jt} = q(v_{jt}, f_{jt}; \theta) + \omega_{jt} + \epsilon_{jt}$$

including measurement error ϵ_{jt}, the relation between measured and true (log) output is $q_{jt} = q^*_{jt} + \epsilon_{jt}$ (or $q^*_{jt} = q_{jt} - \epsilon_{jt}$). Therefore

$$Q^*_{-jt} = \sum_{k \neq j} Q^*_{kt} = \sum_{k \neq j} \exp(q^*_{kt}) = \sum_{k \neq j} \exp(q_{kt} - \epsilon_{kt})$$

and we have[19]

$$V_{jt} = V(\omega_{jt}, F_{jt}, \sum_{k \neq j} \exp(q_{kt} - \epsilon_{kt}), Z_t) \tag{32}$$

or analagously with logged inputs as

$$v_{jt} = v(\omega_{jt}, f_{jt}, \sum_{k \neq j} \exp(q_{kt} - \epsilon_{kt}), z_t) \tag{33}$$

As well known, OP/LP-style approaches require the assumption that the function to be inverted is strictly monotonic in the productivity shock ω_{jt}[20]. We also need to make this strict monotonicity assumption on (33) to utilize it in estimation. This is a substantive assumption. While LP, for example, find relatively simple conditions under which this is the case in a perfectly competitive environment, this situation is different in imperfectly competitive environments. In particular, whether (33) holds depends on the residual demand curves generated by the underlying primitives. It may not hold in some cases - e.g. if the increased production generated by an increased ω_{jt} induces a firm to reduce v_{jt} so as to mitigate price decreases from moving down the demand curve. Biondi (2022) discusses more primitive conditions on demand/residual demand functions such that this strict monotonicity holds.[21]

Under this strict monotonicity assumption one can invert (33) and substitute into the production function

[19] In the logit case we subsume product characteristics X_{jt} into F_{jt} to conserve notation. This implicitly allows the production function to depend on X_{jt}, though if one does not want this to be the case, one can simply impose the restriction that the production function $q\,()$ does not depend the x_{jt} component of f_{jt}.

[20] There is some scope to allow weak monotonicity, see, e.g. LP.

[21] One advantage of the DP and TC approaches is that they do not require such an assumption since there is no first stage inversion necessary.

to get

$$q_{jt} = q(v_{jt}, f_{jt}; \theta) + v^{-1}(v_{jt}, f_{jt}, \sum_{k \neq j} \exp(q_{kt} - \epsilon_{kt}).z_t) + \epsilon_{jt}$$

and since $q(v_{jt}, f_{jt}; \theta)$ cannot generally be separately identified from the non-parametrically treated v^{-1} in this first stage equation, we combine the two functions together (abusing notation by redefining $v^{-1}() = q() + v^{-1}())$ as

$$q_{jt} = v^{-1}(v_{jt}, f_{jt}, \sum_{k \neq j} \exp(q_{kt} - \epsilon_{kt}), z_t) + \epsilon_{jt} \tag{34}$$

Again, the issue here is that since the true quantities $\sum_{k \neq j} \exp(q_{kt} - \epsilon_{kt})$ are not fully observed, straightforward OP/LP-style non-parametric estimation of this equation is not possible.

However, we can leverage the structure in which the additional unobservables ϵ_{kt} enter (34). In particular, consider the set of (34) across *all* firms in market t

$$\begin{aligned} q_{1t} &= v^{-1}(v_{1t}, f_{1t}, \sum_{k \neq 1} \exp(q_{kt} - \epsilon_{kt}), z_t) + \epsilon_{1t} \\ &. \\ q_{Jt} &= v^{-1}(v_{Jt}, f_{Jt}, \sum_{k \neq J} \exp(q_{kt} - \epsilon_{kt}), z_t) + \epsilon_{Jt} \end{aligned} \tag{35}$$

Conditional on observables (q_{jt}, v_{jt}, f_{jt}) for all j and z_t, this is a system of J equations in J unobservables $\epsilon_{1t}, ..., \epsilon_{Jt}$. To make progress here under the classical measurement error assumption that the ϵ_{jt} are mean independent of everything but the q_{jt}'s, we seek assumptions under which (35) can be inverted for $\epsilon_{1t}, ..., \epsilon_{Jt}$ conditional on observables. To do this, consider "reduced form" mappings associated with the above Quantity Setting and Logit Nash Bertrand games, respectively

$$\begin{pmatrix} v_{1t} \\ . \\ v_{Jt} \end{pmatrix} = F_C \left(\begin{pmatrix} f_{1t} \\ . \\ f_{Jt} \end{pmatrix}, \begin{pmatrix} \omega_{1t} \\ . \\ \omega_{Jt} \end{pmatrix}, z_t \right)$$

and

$$\begin{pmatrix} p_{1t} \\ . \\ p_{Jt} \end{pmatrix} = F_L \left(\begin{pmatrix} f_{1t} \\ . \\ f_{Jt} \end{pmatrix}, \begin{pmatrix} x_{1t} \\ . \\ x_{Jt} \end{pmatrix}, \begin{pmatrix} \omega_{1t} \\ . \\ \omega_{Jt} \end{pmatrix}, z_t \right)$$

F_C and F_L map each game's predetermined variables (f_{jt}'s, ω_{jt}'s, z_t, plus x_{jt}'s in logit) into equilibrium choices of variable inputs or prices.

Assumption 5 *The primitives of the model $P()$, $Q()$, and $P_V()$ (or $f()$, $Q()$, and $P_V()$) are such that the mapping F_C (or F_L) is invertible in $(\omega_{1t}, ..., \omega_{Jt})$*

This assumption implies that conditional on fixed inputs, characteristics, and z_t, knowing all firms' observed actions (either variable inputs v's or prices p's) is sufficient to infer the vector of their productivity shocks. This is essentially a multifirm version of the usual single equation strict monotonicity requirement used in OP/LP-style approaches. While again this is a non-trivial assumption for similar reasons as discussed above, there is precedent for related assumptions (see e.g. Petropoulous (2000), ABBP, and Dhyne, Petrin, Smeets, and Warsynski (2022)). The assumption doesn't strictly rule out multiple equilibria, but it is hard to imagine it holding with multiple equilibria unless all markets have same equilibrium selection mechanism. Again there is precedent for such an assumption in the dynamic game literature (e.g. Bajari, Benkard, and Levin (2007)).

Under Assumption (5) we have the following

Theorem 4 *Under Assumption (5), the system of equations defined in (35), i.e.*

$$\begin{aligned} q_{1t} &= v^{-1}(v_{1t}, f_{1t}, \sum_{k \neq 1} \exp(q_{kt} - \epsilon_{kt}), z_t) + \epsilon_{1t} \\ &. \\ q_{Jt} &= v^{-1}(v_{Jt}, f_{Jt}, \sum_{k \neq J} \exp(q_{kt} - \epsilon_{kt}), z_t) + \epsilon_{Jt} \end{aligned} \tag{36}$$

can be inverted for $\epsilon_{1t}, ..., \epsilon_{Jt}$ conditional on observables q_t, v_t, and f_t.

Proof. See Appendix ■

Given this result, one can potentially non-parametrically estimate these first stage equations by, e.g. 1) specifying $v^{-1}()$ as a polynomial or sieve, and 2) searching over the parameters τ of that polynomial or sieve such that the inverted $\epsilon_{1t}, ..., \epsilon_{Jt}$ satisfy the following moment condition

$$E\left[\left(\begin{array}{c} \epsilon_{1t}\left(q_{1t}, .., q_{Jt}, v_{1t}, .., v_{Jt}, f_{1t}, .., f_{Jt}, z_t; \tau\right) \\ . \\ \epsilon_{Jt}\left(q_{1t}, .., q_{Jt}, v_{1t}, .., v_{Jt}, f_{1t}, .., f_{Jt}, z_t; \tau\right) \end{array}\right) \mid v_{1t}, .., v_{Jt}, f_{1t}, .., f_{Jt}, z_t\right] = 0 \tag{37}$$

where $\epsilon_{jt}\left(q_{1t}, .., q_{Jt}, v_{1t}, .., v_{Jt}, f_{1t}, .., f_{Jt}, z_t; \tau\right)$ represents the implied measurement error inverted from (36) given observed data and a candidate value of the parameters τ.[22] This moment condition holds given the assumption that ϵ_{jt} is classical measurement error.[23]

The natural next question is what aspects of the model the set of moments (37) identifies. Given the limited dimension of $v^{-1}()$ relative to the number of conditioning variables in (37) (plus the fact that the same v^{-1} enters all J equations), we suspect that under appropriate regularity conditions, $v^{-1}()$ is identified. However, recall from Section 2 that the goal of the first stage of OP/LP-style approaches is only to identify the measurement errors. Hence, for our purposes we do not need to show that $v^{-1}()$ is identified - as long as the measurement errors $\epsilon_{1t}, ..., \epsilon_{Jt}$ can be identified. This can be shown straightforwardly since we can rewrite (36) as

$$\begin{aligned} q_{1t} - \epsilon_{1t} &= v^{-1}(v_{1t}, f_{1t}, \sum_{k \neq 1} \exp(q_{kt} - \epsilon_{kt}), z_t) \\ &\;\;. \\ q_{Jt} - \epsilon_{Jt} &= v^{-1}(v_{Jt}, f_{Jt}, \sum_{k \neq J} \exp(q_{kt} - \epsilon_{kt}), z_t) \end{aligned} \tag{38}$$

and, given Assumption (5), Theorem (4), and the structure of how q_{jt} and ϵ_{jt} enter, (38) must also be invertible in $q_{1t} - \epsilon_{1t}, ..., q_{Jt} - \epsilon_{Jt}$. Hence, there exists a reduced form

$$\begin{aligned} q_{1t} &= H(v_{1t}, f_{1t}, \boldsymbol{v}_{-1t}, \boldsymbol{f}_{-1t}, z_t) + \epsilon_{1t} \\ &\;\;. \\ q_{Jt} &= H(v_{Jt}, f_{Jt}, \boldsymbol{v}_{-Jt}, \boldsymbol{f}_{-Jt}, z_t) + \epsilon_{Jt} \end{aligned} \tag{39}$$

[22] As this non-linear inversion may be non-trivial to implement in practice, we briefly mention a potential alternative that can potentially eliminate the ϵ_{kt}'s inside the $v^{-1}()$ function on the right hand side of (36). For this we need the researcher to also have access to firm-level revenue data, R_{jt}, that, unlike Q_{jt}, is not contaminated by measurement error (e.g. suppose that mismeasured prices are the reason that Q_{jt} has measurement error). Then in the quantity setting model, under appropriate conditions on demand curves and equilibria (e.g. that all equilibria are on elastic parts of demand curves), there is a one to one mapping between Q's and R's, and the Q_{-jt} term inside inside v^{-1} can be replaced by R_{-jt} . With no measurement error inside v^{-1}, estimation can proceed, e.g. with OLS using a polynomial approximation to $v^{-1}())$.

[23] To clarify given the fact that this is a system of equations, note that when we say classical measurement error we mean we are assuming that the measurement error in q_{jt} is not only mean independent of v_{jt} and f_{jt}, but also of v_{-jt} and f_{-jt}. Note that we do not assume that the measurement error is uncorrelated across j.

Given the classical measurement assumption that the errors are mean independent of all the arguments of H, this is a standard non-parametric problem and we have

Theorem 5 *Under Assumption (5), with sufficient variation in* $(v_{1t}, .., v_{Jt}, f_{1t}, .., f_{Jt}, z_t)$ $H()$ *is non-parametrically identified and the measurement errors* $\epsilon_{1t}, ..., \epsilon_{Jt}$ *are identified.*

Proof. See Matzkin (1994, 2013) ∎

The assumption of "sufficient variation" is not precise and less than ideal, but not surprising given the goal of estimating the production function without having to fully specific other components of firms' decision problems. For example, variation in $f_{1t}, .., f_{Jt}$ will likely depend on the exact nature of firms dynamic investment decisions of, e.g. capital, other fixed factors, and product characteristics - we have left these components of the model unspecified. Variation in $v_{1t}, .., v_{Jt}$ will depend on the entire distribution of productivity shocks $\omega_{1t}, .., \omega_{Jt}$ as well as the exact nature of demand $P()$ in Cournot (or $f()$ in Logit) - again our intent is not require specification of these objects.

4 Second Stage Identification and Estimation

The prior section shows that we can obtain a low-dimensional OP/LP-style first stage inversion in a set of models with imperfect competition and strategic interactions, without fully specifying demand and in some cases the nature of competition (e.g. Cournot vs. collusion). This allows us to recover estimates $\widehat{\epsilon}_{jt}$'s. We now move to the second stage of these procedures, which use moments of the general form $\mathbb{E}\left[\xi_{jt} \mid \mathcal{I}_{jt-1}\right] = 0$ (recall ξ_{jt} is the period t innovation in the productivity shock ω_{jt}) to identify the structural parameters of the production function and productivity shock process.

As discussed in Section 2, the OP/LP-style first stage described in the prior section is just one of at least three different ways that one can get to this point. Alternatively, one could use either the TC or DP approach. Each of these relies on an additional restriction - in TC an additional variable input that enters the production function in Leontief fashion, a in DP a linear productivity shock process - but on the other hand, neither requires the OP/LP-style first stage inversion and associated assumptions. So, for example, the TC or DP approaches might be useful if either 1) is unwilling to assume that the model of imperfect competition is, as per Section 3, homogeneous good, logit, or nested logit such that a low dimension sufficient statistic exists[24], or 2) believes there are other unobservables that make the scalar unobservable assumption

[24] There might be scope for extension to more general models. For example, in a random coefficients model, one could consider

required for the OP/LP-style first stage untenable.

Our main point in this section is to show how in *all three cases*, the nature of imperfectly competitive markets can actually help with respect to some of the identification issues that have been discussed in the production function literature. This is because we can leverage what we term *oligopoly instruments* based on the fact that fixed inputs and productivity shocks of competitors in a given market t determine an individual producer's variable input choice through strategic interactions. In short, the existence of oligopoly interactions means there are additional elements in $\mathcal{I}_{jt-1}$ that can aid with identification. In the next subsections, we 1) show how oligopoly instruments can resolve some existing identification problems, 2) consider properties of optimal oligopoly instruments, and 3) consider use of oligopoly instruments when it is hard to define which firms compete with one another.

4.1 Oligopoly Instruments to Resolve Identification Problems

For simplicity, we illustrate with the two approaches (OP/LP-style and TC) where $\widehat{\epsilon}_{jt}$ has already been identified. As illustrated in Section 2, this means that given observables (and $\widehat{\epsilon}_{jt}$) we can infer ω_{jt}'s given parameters θ of the production function, i.e.

$$\omega_{jt}(\theta) = (q_{jt} - \widehat{\epsilon}_{jt}) - q(v_{jt}, f_{jt}; \theta)$$

and given parameters ρ of the productivity process $\omega_{jt} = g(\omega_{jt-1}, \rho) + \xi_{jt}$,[25] we can additionally infer ξ_{jt}'s, i.e.

$$\xi_{jt}(\theta) = \omega_{jt}(\theta) - g(\omega_{jt-1}(\theta), \rho)$$

This allows us to easily construct an empirical analogue of the moment condition $\mathbb{E}\left[\xi_{jt} \mid \mathcal{I}_{jt-1}\right] = 0$. In contrast, the DP approach (since $\widehat{\epsilon}_{jt}$'s have not been estimated previously) requires one to utilize slightly different moment conditions that also include the unobservables ϵ_{jt} (see (8)), but the logic of the rest of this section still applies. To conserve notation, we redefine $q_{jt} = q_{jt} - \widehat{\epsilon}_{jt}$ for the rest of this section, i.e. q_{jt}

approximating the $(\boldsymbol{\omega}_{-jt}, \boldsymbol{F}_{-jt}, \boldsymbol{X}_{-jt})$ elements of (14) with moments of these distibutions (e.g. the means/variances/covariances of these variables across competitors) to reduce dimensionality. However, one would need to assure the approximation error disappears asymptotically for consistency, and this would still not allow unobserved product characteristics as common in the demand literature (under the presumption of this paper that we dont have the data to estimate demand and "observe" these unobserved characteristics).

[25] We could extend our results to allow the "endogenous" productivity shock process of Doraszelski and Jaumandreu (2013), i.e. $\omega_{jt} = g(\omega_{jt-1}, r_{it}, \rho) + \xi_{jt}$, for some observed r_{it} (e.g. R&D expenditures).

represents the originally measured q purged of the estimated measurement error $\widehat{\epsilon}_{jt}$.[26]

The typical timing and information set assumptions of the model imply the following conditional moment restriction

$$\mathbb{E}\left[\xi_{jt}(\theta) \mid \mathcal{I}_{jt-1}\right] = \mathbb{E}\left[\omega_{jt}(\theta) - g\left(\omega_{jt-1}(\theta); \rho\right) \mid \mathcal{I}_{jt-1}\right] = 0 \qquad \text{at } \theta = \theta_0 \tag{40}$$

at the true parameter values, which in turn implies the unconditional moment restriction

$$\mathbb{E}\left[\left(\omega_{jt}(\theta) - g_t\left(\omega_{jt-1}(\theta); \rho\right)\right) \otimes \mathcal{I}_{jt-1}\right] = 0 \qquad \text{at } \theta = \theta_0 \tag{41}$$

If, as typical, f_{jt} is assumed chosen at $t-1$, v_{jt} is assumed chosen at t, and ω_{jt} assumed observed at t, then there are many elements of $\mathcal{I}_{jt-1}$. $\mathcal{I}_{jt-1}$ includes f_{jt}, v_{jt-1}, q_{jt-1}, and ω_{jt-1}, and all lags of these variables (as well as all measurable functions of all these variables). An important contribution of this paper is to note that in the context of oligopolistic competition models like those considered here, $\mathcal{I}_{jt-1}$ also can contain these variables for *competing* firms, e.g. $\boldsymbol{f}_{-jt}$, $\boldsymbol{v}_{-jt-1}$, $\boldsymbol{q}_{-jt-1}$, and $\boldsymbol{\omega}_{-jt-1}$ and lags of these variables. We will momentarily argue that this has important benefits for estimation based on (41), but before doing this we first discuss three relevant preliminary points.

First, note that for convenience in the below we will assume that competing firms' $\boldsymbol{f}_{-jt}$, $\boldsymbol{v}_{-jt-1}$, $\boldsymbol{q}_{-jt-1}$, and $\boldsymbol{\omega}_{-jt-1}$ are in fact in $\mathcal{I}_{jt-1}$. However, with additional notation this could easily be relaxed to the case where firm j only observe some information about its competitors variables. It is important to note that doing this in the context of an OP/LP-style first stage inversion would require extending the models of Section 2 to ones of imperfect information. And clearly if firm j observes *no* information about its competitors variables, the oligopoly instruments we consider will not be useful. Following the suggestions in Ackerberg (2023), one could potentially test different timing and information set assumptions, e.g. the additional assumption that ω_{-jt} is in $\mathcal{I}_{jt-1}$, but we do not investigate this here.

Second, note that while one can obtain efficiency by enforcing (41) using the entire $\mathcal{I}_{jt-1}$ (to the extent it is observed by the econometrician) and weighting the moment conditions appropriately, this is often not done in practice. This is because with so many elements of $\mathcal{I}_{jt-1}$ constructing optimal weights tends to generate small sample biases, i.e. the "many instruments" problem described in, e.g. Bound, Jaeger, and

[26] This highlights that a third way to avoid an OP/LP-style inversion is to simply start with a model with no measurement error, i.e. assume $\epsilon_{it} \equiv 0$. If one is willing to make this additional assumption, then one can directly move to these second stage moment conditions without a first stage inversion.

Baker (1996). Instead, researchers will often use an unconditional moment condition of the form

$$\mathbb{E}\left[\left(\omega_{jt}(\theta) - g_t\left(\omega_{jt-1}(\theta);\rho\right)\right) \otimes m(\mathcal{I}_{jt-1})\right] = 0 \tag{42}$$

where m is of reduced dimension. For example, papers based on OP, LP, and ACF with Cobb-Douglas production functions will often use just f_{jt}, v_{jt-1}, and q_{jt-1} (or ω_{jt-1} defined as a function of the parameters) as instruments, resulting in an exactly identified model and no need for a weight matrix.

Third, there has been much recent discussion regarding conditions under which the moments (42) (or (41)) identify the production function parameters. In particular, in the case of perfect competition, GNR derive an important result that if there is no across-firm heterogeneity in prices of variable (and assumed static) inputs v_{jt} or demand, then these moments do not identify a non-parametric production function $q_{jt} = q_t(v_{jt}, f_{jt}) + \omega_{jt}$. Note that this result is not dependent on the choice of $m()$, i.e. including further lags of f_{jt}, v_{jt-1}, q_{jt-1}, and ω_{jt-1} does not help. The intuition is that in such a model, optimal v_{jt} is a function of only f_{jt} and ω_{jt}. This functional dependence issue means that in this model, v_{jt-1} (nor further lags) is not useful as an instrument to identify the production function.

There are a several existing ways of avoiding the GNR non-identification result. DLW, for example, argue that observing additional, firm-specific factors that determine firms' variable input choice v_{jt} is sufficient for identification (assuming ξ_{jt} is mean independent of these observables). These could include observed input or output market shifters, or v_{jt-1} if firms' choices of variable input have dynamic implications. The appendix of LP (parametrically) and GNR (non-parametrically) respectively, add restrictions from first order conditions when firms are in perfectly competitive environments to generate identification.[27] ACF suggests that identification can be obtained if there are serially correlated firm-specific shocks to prices of v_{jt}, even if they are unobserved to the econometrician - however, this approach does not generally work in the context of an OP/LP-style inversion of v_{jt} (because it violates the scalar unobservable assumption). Flynn, Gandhi, and Traina (2021) (also see DLS) use the strategy of assuming constant returns to scale - which implies there is no need to estimate a coefficient on v_{jt}.

The main point of this section of the paper is that oligopoly instruments provide an *additional way* to avoid the GNR non-identification result.. In particular, as discussed above, in an oligopoly setting v_{jt} is not only a function of f_{jt} and ω_{jt}, but also generally of $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt}$. This can alleviate the functional dependence issue described by GNR. And while $\boldsymbol{\omega}_{-jt}$ is not in $\mathcal{I}_{jt-1}$, $\boldsymbol{f}_{-jt}$ is, and so is $\boldsymbol{\omega}_{-jt-1}$, which

[27] More recent papers further developing these methods include Navarro and Rivers (2018) and Pan (2022).

is generally correlated with $\boldsymbol{\omega}_{-jt}$ (unless the ω process is not serially correlated). Hence, in oligopolistic settings, $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt-1}$ can be useful "instruments" for identifying the production function parameters.

Note the intuition behind this idea. Since the variable input v_{jt} is endogenous (i.e. correlated with $\xi_{jt} = \omega_{jt} - g_t(\omega_{jt-1})$), identifying its effect on output requires some source of exogenous variation. We are simply observing that in an oligopoly situation natural instruments for v_{jt} are competitors' $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt-1}$. Both are orthogonal to ξ_{jt} based on the timing and information set assumptions of the model, so they are valid instruments. As for the relevance condition, $\boldsymbol{f}_{-jt}$ directly impacts v_{jt} through the oligopolistic structure of the model, and while v_{jt} directly depends on $\boldsymbol{\omega}_{-jt}$ (rather than $\boldsymbol{\omega}_{-jt-1}$), $\boldsymbol{\omega}_{-jt-1}$ is generally correlated with $\boldsymbol{\omega}_{-jt}$, making it also a potentially relevant instrument. This can be interpreted as a supply side application of BLP "competitive" instruments for demand estimation, i.e. where aspects of competitors (in that case typically competitor product characteristics) are used to instrument for endogenous price.

4.2 Approximations to Optimal Instruments

Given that these competitive instruments can help one avoid the production function non-identification result of GNR, we now study the choice of $m(\mathcal{I}_{jt-1})$ in (42). As noted above, given the size of $\mathcal{I}_{jt-1}$, small-sample biases are a significant concern with the "brute force" method of using all functions of all of the observed elements of $\mathcal{I}_{jt-1}$ as instruments (or a large number of functions of those elements). These biases can come from estimating a large weight matrix for these moments in order to weight them efficiently. An alternative is to construct "optimal" instruments associated with the conditional moment restriction, i.e. following Chamberlain (1987). This has the advantage of potentially reducing the number of instruments by enforcing economic restrictions implied by the production function (and possibly the nature of competition).

In our oligopolistic situation, fully characterizing the optimal instruments is challenging. As detailed below, it requires computing expectations of period t decisions conditional on $\mathcal{I}_{jt-1}$, so it depends on structural objects such as the demand curve, the precise contents of the information set $\mathcal{I}_{jt-1}$, and the entire distributions of unobservable terms - objects that have thus far not been specified. We take an alternative approach where we consider additional economic assumptions that either 1) simplify computation of the optimal instruments, or 2) reduce the dimension of the $m(\mathcal{I}_{jt-1})$ that needs to be considered for optimality. These additional assumptions include things like a Cobb-Douglas assumption, AR(1) productivity, and linear demand. It is important to note that these additional assumptions are not required for consistency of the estimates, i.e. even if these additional assumptions are incorrect, the parameter estimates are still consistent (assuming identification conditions hold). The tradeoff is instead one between not quite achieving full efficiency (to the extent the assumptions are incorrect) versus reducing small sample biases (because the assumptions allow us to reduce the dimension of $m(\mathcal{I}_{jt-1})$ based on theory). This type of approach - i.e. imposing additional assumptions to simplify and/or approximate optimal instruments - has also been taken in the context of discrete choice demand estimation, e.g. Berry, Levinsohn, and Pakes (1995), Reynaert and Verboven (2015) and Gandhi and Houde (2021).

Also similar to the literature on optimal instruments for demand models, our simplifying assumptions have economic intuition. First, to clarify notation, denote the level of fixed inputs and productivity for a *single* competitor of firm j in market t as f_{kt} and ω_{kt} (this contrasts with $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt}$, which represent vectors of these objects for *all* competitors of firm j in market t). Our first result formalizes the idea that since in some cases f_{kt} and ω_{kt} enter firm k's production function in the same way, j's optimal choice of variable input will only depend on f_{kt} and ω_{kt-1} through a scalar summary statistic. In other words, we develop assumptions under which we can reduce the dimensionality of the instrument set coming from each of j's individual competitors k. Our second result reduces the dimensionality of the instrument set *across* j's competitors. Specifically, we construct assumptions under which the fixed inputs and productivity shocks af all j's competitors, $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt}$ affect the optimal instrument through only a scalar. Loosely speaking, the intuition here is related to the homogenous good quantity setting result that competitors matter for j's FOC only through the sum of their quantities. This also means that the argument is similar to that in Section 2 - however, additional assumptions are needed.

4.3 Within Firm

Here we focus on reducing the dimensionality of the instruments by constructing assumptions under which, for a specific competitor k, the instruments f_{kt} and ω_{kt-1} can be combined into a single instrument. Specifically, we show that under these assumptions, instead of using f_{kt} and ω_{kt-1} as individual instruments (for firm j's moment condition), we can alternatively use the linear combination $\theta_f f_{kt}+ \rho\omega_{kt-1}$, without sacrificing efficiency. Again, this alleviates having to use the data to individually weight separate moment conditions corresponding to f_{kt} and ω_{kt-1} (and the potential small sample biases associated with doing that). There is some clear economic intuition here - in many competitive models, the effect of firm B's cost technology on firm A is only through firm B's marginal cost curve. So it is possible that multiple aspects of firm B's technology affect firm A through a scalar. Essentially this section posits specific assumptions under which this is the case.

We do this under the assumption of Cournot competition with a Cobb Douglas production function and an AR(1) Markov process. We continue to assume only a single variable input and assume a single competitor k and two periods of data to illustrate - we discuss generalizations momentarily. Under these conditions, with 0 indexing the first period of data and 1 indexing the second, the conditional moment condition is

$$\begin{aligned} & \mathbb{E}\left[(q_{j1} - \theta_0 - \theta_v v_{j1} - \theta_f f_{j1}) - \rho\left(q_{j0} - \theta_0 - \theta_v v_{j0} - \theta_f f_{j0}\right) \mid \mathcal{I}_{j0}\right] && (43) \\ = \ & \mathbb{E}\left[q_{j1} - \rho q_{j0}\right) - \theta_0\left(1-\rho\right) - \theta_v\left(v_{j1} - \rho v_{j0}\right) - \theta_f\left(f_{j1} - \rho f_{j0}\right) \mid \mathcal{I}_{j0}] && (44) \end{aligned}$$

As shown by Chamberlain (1987), optimal instruments $m(\mathcal{I}_{jt-1})$ for the corresponding unconditional moment (as in (42)) are constructed by taking derivatives of the conditional moment condition with respect to the parameters. For example, for the parameter θ_f, this derivative is

$$\mathbb{E}\left[-\left(f_{j1} - \rho f_{j0}\right) \mid \mathcal{I}_{j0})\right] = -\left(f_{j1} - \rho f_{j0}\right)$$

i.e., since $f_{j0}, f_{j1} \in \mathcal{I}_{j0}$ by the timing and information set assumptions, the optimal instrument corresponding to θ_f is simply $f_{j1} - \rho f_{j0}$. For the parameter ρ, the derivative is

$$\mathbb{E}\left[-\left(q_{j0} - \theta_0 - \theta_v v_{j0} - \theta_f f_{j0}\right) \mid \mathcal{I}_{j0}\right] = -\left(q_{j0} - \theta_0 - \theta_v v_{j0} - \theta_f f_{j0}\right)$$

i.e., again, because $q_{j0}, v_{j0}, f_{j0} \in \mathcal{I}_{j0}$, the corresponding optimal instrument is $(q_{j0} - \theta_0 - \theta_v v_{j0} - \theta_f f_{j0})$, i.e. the "constructed" value of ω_{j0} given a set of parameters.

On the other hand, because v is endogenous, the optimal instrument for the parameter θ_{v1} is not so simple. Chamberlain's formula gives

$$\mathbb{E}\left[-\left(v_{j1} - \rho v_{j0}\right) \mid \mathcal{I}_{j0}\right] = -\mathbb{E}\left[v_{j1} \middle| \mathcal{I}_{j0}\right] + \rho v_{j0} \tag{45}$$

but since $v_{j1} \notin \mathcal{I}_{j0}$, this does not simplify. Moreover, as noted above, the conditional expectation $\mathbb{E}\left[v_{j1} \middle| \mathcal{I}_{j0}\right]$ is complicated - depending on many parts of the model that we have not specified, e.g. the demand curve and the full distribution of ω.

That said, under the above assumptions, $\mathbb{E}\left[v_{j1} \middle| \mathcal{I}_{j0}\right]$ has simpler structure. Given the results in Sections

2 and 3, we can write the equilibrium choice of v_{j1} as

$$v_{j1} = v(\omega_{j1}, f_{j1}, \omega_{k1}, f_{k1}, z_1), \tag{46}$$

where again z_1 represent market level variables affecting demand or costs (e.g. demand or input price shifters). The elements of ω_{k1} and f_{k1} enter the system of first order conditions only through the firms' respective production functions, which under the Cobb Douglas and AR(1) assumptions can be written as

$$\begin{aligned} Q_k(v_{k1}, \omega_{k1}, f_{k1}) &= \exp\left(\theta_0 + \theta_f f_{k1} + \theta_v v_{k1} + \omega_{k1} + \epsilon_{k1}\right) \\ &= \exp\left(\theta_0 + \theta_f f_{k1} + \theta_v v_{k1} + \rho\omega_{k0} + \xi_{k1} + \epsilon_{k1}\right) \end{aligned}$$

This means that we can alternatively write (46) as

$$v_{j1} = v(\omega_{j1}, f_{j1}, \theta_f f_{k1} + \rho\omega_{k0}, \xi_{k1}, z_1), \tag{47}$$

where the vectors of competing firms fixed inputs and prior productivity shocks, f_{k1} and ω_{k0} (both in $\mathcal{I}_{j0}$) enter *only* through their sum $\theta_f f_{k1} + \rho\omega_{k0}$. However, this does not neccesarily imply that $\mathbb{E}\left[v_{j1}|\, \mathcal{I}_{j0}\right]$ only depends on this sum. For example, we have not specified what the firm knows about z_1 at $t = 0$. It is possible that at that point in time the firm might know the entirety of z_1 (i.e. $z_1 \in \mathcal{I}_{j0}$); it might know nothing about z_1 apart from its distribution; or, it might "know" part of z_1 (e.g. it observes partially informative signals about z_1, or it knows some elements with certainty, but not other elements). Similarly, we have assumed ω_{j1} and ξ_{k1} are not in $\mathcal{I}_{j0}$ but we have not specified their full distribution. It is possible that $\mathbb{E}\left[v_{j1}|\, \mathcal{I}_{j0}\right]$ depends on the values of f_{k1} and ω_{k0} individually through their containing information on the distributions of z_1, ω_{j1} and ξ_{-j1}. For example, if z_1 is not in $\mathcal{I}_{j0}$, and the distribution of z_1 given $\mathcal{I}_{j0}$ for some reason depends on the individual elements of f_{k1} and ω_{k0} (rather than just the sum $\theta_f f_{k1} + \rho\omega_{k0}$), then $\mathbb{E}\left[v_{j1}|\, \mathcal{I}_{j0}\right]$ will depend on the values of the individual elements. The same holds for ω_{j1} and $\boldsymbol{\xi}_{-j1}$ (note that while the model does make assumptions on the means of these variables, their higher order moments could potenially depend on the values of f_{k1} and/or ω_{k0} individually).

Hence, we need an additional assumption to rule out this "indirect" effect, specifically

Assumption 6 *The distribution of* $p(\omega_{j1}, \xi_{k1}, z_1 \mid \mathcal{I}_{j0})$ *is such that*

$$p(\omega_{j1}, \xi_{-j1}, z_1 \mid \mathcal{I}_{j0} \backslash (f_{k1}, \omega_{k0}) = \Psi_1,\ f_{k1} = \Psi_2, \omega_{k0} = \Psi_3)$$
$$= p(\omega_{j1}, \xi_{-j1}, z_1 \mid \mathcal{I}_{j0} \backslash (f_{k1}, \omega_{k0}) = \Psi_1,\ \theta_f f_{k1} + \rho\omega_{k0} = \theta_f \Psi_2 + \rho\Psi_3)$$

This assumption rules out f_{k1} and ω_{k0} affecting the joint distribution of ω_{j1}, ξ_{k1}, and z_1 other than through the sum $\theta_f f_{k1} + \rho\omega_{k0}$. Now we have

Theorem 6 *In the above model, under Assumption (6), the optimal instruments* $m^*(\mathcal{I}_{j0})$ *depend on* f_{k1} *and* ω_{k0} *only through the sum* $\theta_f f_{k1} + \rho\omega_{k0}$.

Proof. As detailed above, in this model, under Assumption (6), $\mathbb{E}[v_{j1} | \mathcal{I}_{j0}]$ depends on f_{k1} and ω_{k0} only through the sum $\theta_f f_{k1} + \rho\omega_{k0}$ ■

The simple intuition behind this Theorem suggests it can be generalized in a number of directions, e.g. $t > 2$, $J > 2$, a more general Markov structure $g_t()$ on the evolution of the productivity shock, or a more general log production function that is addively separable in f_{jt}, i.e. $q(v_{jt}, f_{jt}; \theta) + \omega_{jt} = q_v(v_{jt}; \theta) + q(f_{jt}; \theta) + \omega_{jt}$ (clearly if f interacts with v independently of ω, it will not hold). It also generalizes to the logit and nested logit models, but only for elements of f_{k1} that are not also product characteristics entering the demand side of the model. Regardless, note that regardless of the technological structure on $q()$ and $g_t()$, these generalizations will always require an additional assumption analagous to Assumption (6).[28]

Lastly, note that Theorem (6) does not imply that $\theta_f f_{k1} + \rho\omega_{k0}$ itself is an optimal instrument. In general, the optimal instrument will be some *function* of this sum, and this function will depend on aspects of the model that have been left unspecified, e.g. the shape of demand. Hence, to the extent one wants to fully leverage optimal instruments, one would want to follow, e.g., Berry, Levinsohn, and Pakes (1995) and use a polynomial in $\theta_f f_{k1} + \rho\omega_{k0}$ as instruments. This can also result in a proliferation of moment conditions, but less so that if one did not make use of these theoretical restrictrictions and used polynomials in f_{k1} and ω_{k0} individually. We examine some of these issues in our Monte-Carlo experiments.

4.4 Across Firms

Next we ask the question whether instruments can be aggregated across competitors. In other words, suppose that j faces two competitors, k and l. Under the assumptions in the prior section, we know that

[28] Note that if one were willing to make the additional assumption that innovations in the productivity process are mean independent across firms, then one could use $\theta_f f_{k1} + \omega_{k1}$ instead of $\theta_f f_{k1} + \rho\omega_{k0}$.

the optimal instruments depend only on the two scalars $\theta_f f_{k1} + \rho\omega_{k0}$ and $\theta_f f_{l1} + \rho\omega_{l0}$, but we ask whether there are additional assumptions such that these can be combined into some *scalar* instrument

$$m(\theta_f f_{k1} + \rho\omega_{k0}, \theta_f f_{l1} + \rho\omega_{l0}; \theta, \rho)$$

that does not depend on any unknown parameters (e.g. aspects of unspecified demand). It turns out we can, but this requires considerably more assumptions than in the prior section, even with homogeneous goods. The intuition is that while, e.g. in a Cournot model, firm j only cares about the sum of quantities of its competitors, these quantities are determined in equilibrium, so the aggregation of the factors determining these quantities (e.g. $\theta_f f_{k1} + \rho\omega_{k0}$ and $\theta_f f_{l1} + \rho\omega_{l0}$) is non trivial. We again emphasize that the additional assumptions are only necessary for the efficiency property to hold, i.e. their violation does not affect consistency.

We maintain Assumption (6) (expanded to multiple competitors), and additionally assume

Assumption 7 *Production technology is Cobb-Douglas with constant returns to scale in the variable input, i.e.*

$$q_{jt} = \theta_0 + v_{jt} + \theta_f f_{jt} + \omega_{jt}$$

Assumption 8 *Linear aggregate demand*

$$P_t = \alpha - \beta Q_t$$

Assumption 9 *Given the predetermined variables (f_{jt}'s, ω_{jt}'s, and z_t) firms simultaneously choose v_{jt} to maximize profits under a non-common knowledge belief structure in which each firm j believes that each other firm k believes that all other firms have $f = f_{kt}$ and $\omega = \omega_{kt}$.*

Assumptions (7) and (8) are self explanatory. Assumption (9) is a behavioral assumption that is more unusual. It can be summarized by "everybody thinks everybody (else) thinks everybody is like them" (ETETEILT) - i.e. each firm knows that there is heterogeneity in f_{jt} and ω_{jt} across j, and knows the values of that heterogeneity for all other firms, but believes that all other firms k play more naively and assume that everybody is like themselves (i.e. like k). In other words, firm j believes that all other firms k believe they are playing a *homogeneous firm* Cournot game where everybody has $f = f_{kt}$ and $\omega = \omega_{kt}$. This is somewhat reminiscent of level-k rationality (e.g. Nagel (1995)) in that players believe other players are not as sophisticated as them, but distinct as the beliefs are different.[29]

Assumption (9) is helpful in our situation because it simplifies how the heterogeneity enters the first order conditions. We know that firm j only cares about the total quantity of other firms, $\sum_{k\neq j} Q(V_{kt}, F_{kt};\theta)\exp(\omega_{jt})$. With ETETEILT, firm j believes each element in this sum is the solution to a *homogeneous firm* Cournot game with costs based on f_{kt} and ω_{kt}. Thus, each element of the sum only depends on f_{kt} and ω_{kt}. This is much simpler than Cournot games with heterogeneous firms, which typically do not have closed form solutions, and with the additional assumptions of CRS in the variable input and linear demand, we obtain

Theorem 7 *Under Assumptions (6), (7), (8), and (9), firm j's optimal variable input choice* $v_{jt} = v(\omega_{jt}, f_{jt}, \boldsymbol{\omega}_{-jt}, \boldsymbol{f}_{-jt}, z_t)$ *depends on the vectors* $\boldsymbol{\omega}_{-jt}$ *and* $\boldsymbol{f}_{-jt}$ *only through their dimension* $J-1$ *and the scalar*

$$\sum_{k\neq j}\frac{1}{\exp(\theta_0+\theta_f f_{kt}+\omega_{kt})}$$

Proof. See Appendix ■

Since the scalar summary statistic in Theorem (7) does not depend on parameters other than those of the structural objects being estimated, it can serve as the basis of an optimal instrument calculation similar to the prior section. To make things most straightforward, suppose that one is willing to strengthen the assumption $\mathbb{E}\left[\xi_{jt} \mid \mathcal{I}_{jt-1}\right]=0$ to $\mathbb{E}\left[\xi_{jt} \mid \mathcal{I}_{jt-1}, \boldsymbol{\omega}_{-jt}\right]=0$ - this is essentially making the additional assumption that the ξ_{jt} are mean independent across firms. Based on Theorem (7), this means that the optimal instrument for $\mathbb{E}\left[v_{jt}\middle| \mathcal{I}_{jt-1}, \boldsymbol{\omega}_{-jt}\right]$ will depend on $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt}$ only through the scalar $\sum_{k\neq j}\frac{1}{\exp(\theta_0+\theta_f f_{kt}+\omega_{kt})}$.[30] This is a potentially large dimensionality reduction, especially if J is large and there are multiple fixed inputs. Again, one would include a polynomial in this scalar instrument to fully leverage its optimality. If one is not willing to make the additional assumption that $\mathbb{E}\left[\xi_{jt} \mid \boldsymbol{\omega}_{-jt}\right]=0$, an alternative is to use $\sum_{k\neq j}\frac{1}{\exp(\theta_0+\theta_f f_{kt}+\rho\omega_{kt-1})}$, but since the innovation component of ω_{kt} (i.e. ξ_{kt}) is in the denominator, this is only an approximation as the variance of the innovation term goes to zero.

While the above is one strategy to limit the proliferation of instruments across firms, there are others. For example, instead of a basis of polynomials in $f_{jt}, \omega_{jt-1}, \boldsymbol{f}_{-jt}, \boldsymbol{\omega}_{-jt-1}$, and z_t one could limit dimensionality

[29] Like level-k rationality, ETEETEILT can be iterated, e.g. "everybody thinks everybody (else) thinks everybody (else) thinks everybody is like them"

[30] And if J varies across markets, the number of competitors $J-1$. However, an observation from the proof of Theorem (7) is that if one adds the additional assumption that the inverse demand intercept α scales in $J+1$ (i.e. that $\frac{\alpha}{J+1}$ is constant, consistent with more firms entering stronger markets), then one no longer needs to add $J-1$ if J varies across markets. In this case the optimal instrument will depend on $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt}$ only through $\frac{1}{J+1}\sum_{k\neq j}\frac{1}{\exp(\theta_0+\theta_f f_{kt}+\omega_{kt})}$.

(and enforce exchangeability in $-j$) by using polynomials in $f_{jt}, \omega_{jt-1}, M(\boldsymbol{f}_{-jt}, \boldsymbol{\omega}_{-jt-1})$, J, and z_t, where M includes moments of its arguments, e.g. $\text{mean}(\boldsymbol{f}_{-jt})$, $\text{var}(\boldsymbol{f}_{-jt})$, $\text{cov}(\boldsymbol{f}_{-jt}, \boldsymbol{\omega}_{-jt-1})$, and higher order moments (under the assumptions of the prior section, one could instead use $M(\theta_f \boldsymbol{f}_{-jt} + \rho \boldsymbol{\omega}_{-jt-1})$), describing the nature of j's competition. Of course, if one uses more than first moments, this approach will require more instruments than if one calculates instruments based on the assumptions described above.

4.5 Unobserved Competitors

Using the oligopoly instruments as described in the previous subsections requires observing all the competitors of firm j. In some datasets, this may not be available. For example, a dataset may include only a sample of firms so that the researcher does not observe all (or any) competitors of j. Or, a dataset may include all firms, but the researcher is not sure exactly who competes with who - e.g. if firms differ based on geographic location, are all firms within 10 miles competitors, or are all firms within 100 miles competitors?

An interesting question is whether the idea behind oligopoly instruments can still be helpful here. We consider this under the polar opposite case than above, i.e. we know the firms in our dataset have competitors, but instead of observing all of the competitors we *do not observe any of these competitors.*[31] There are two immediate implications of this. First, since firms do have competitors but they are unobserved, OP/LP-style first stages violate the scalar unobservable assumption w.r.t. $v()$ and cannot be utilized. However, one can still proceed with the TC or DP approaches, which do not require this scalar unobservable assumption. Second, without observing anything about the competitors of each j, one obviously cannot construct any of the instruments described above for use in the second stage.

On the other hand, even if the econometrician does not observe the competitive factors in $\mathcal{I}_{jt-1}$ required to explicitly form the oligopoly instruments, it is possible that there are other elements of $\mathcal{I}_{jt-1}$ that are observed to the econometrician and *correlated* with the oligopoly instruments that can generate identification. In particular, we focus on the lagged value of j's variable input, i.e. v_{jt-1}. By the timing and information assumptions of the model, $v_{jt-1} \in \mathcal{I}_{jt-1}$. Moreover, in the oligopoly models above, v_{jt-1} is chosen in equilibrium as a function of the (now unobserved) $\boldsymbol{f}_{-jt-1}$ and $\boldsymbol{\omega}_{-jt-1}$. Lastly, if fixed factors f are subject to dynamic accumulation and there is any persistence in ω, then $\boldsymbol{f}_{-jt-1}$ and $\boldsymbol{\omega}_{-jt-1}$ will be correlated with $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt}$. Hence, v_{jt-1} may be correlated with the oligoply instruments and its use as an instrument may provide identification. We verify this in a simple example in our Monte-Carlo experiments below.

[31] One could use a combination of the techniques in intermediate cases.

Using v_{jt-1} as an informative instrument is not a new idea. It has been used as an instrument many times in OP/LP-style, TC, and DP approaches. But it has been subject to criticism, (appropriately so) because of the GNR result that it is an uninformative instrument in a perfectly competitive model (with no across firm input price variation), and hence in that case should not identify the production function. Our argument here is that oligopoly situations provide an additional rationale for v_{jt-1} to be a useful instrument for estimating production functions. Of course, it is not the ideal oligopoly instrument - as per the previous discussion, $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt-1}$ (or $\boldsymbol{\omega}_{-jt}$) would be. v_{jt-1} as an instrument only leverages this variation in competition indirectly - for example it only depends on $\boldsymbol{f}_{-jt-1}$ (not $\boldsymbol{f}_{-jt}$) - and its strength will depend on, e.g. the persistence in $\boldsymbol{f}_{-jt-1}$ and $\boldsymbol{\omega}_{-jt-1}$.[32] This is something else that we evaluate in our Monte-Carlo experiments.

5 Monte Carlo Experiments

Our Monte-Carlo experiments examine the use of oligopoly instruments. To keep things as simple as possible, we do this in a two period Cournot duopoly model with one fixed input, and additionally make the assumption that the econometrician observes ω_{j0}. We take a TC approach, which means that we can eliminate ϵ_{jt} from the model without needing an OP/LP-style inversion (which is important because we want to consider the possibility that competitors are not observed). The setup then becomes very simple as the second stage estimating equation is

$$q_{j1} = \theta_0 + \theta_f f_{j1} + \theta_v v_{j1} + \rho\omega_{j0} + \xi_{j1}$$

with the moment condition

$$E\left[\xi_{jt}|I_{jt-1}\right] = E\left[q_{j1} - \theta_0 - \theta_f f_{j1} - \theta_v v_{j1} - \rho\omega_{j0}|I_{j0}\right] = 0 \quad \text{ at } \theta = \theta_0$$

While treating ω_{j0} as observed is unrealistic (normally it is a function of parameters and lagged variables), it is helpful because it 1) avoids complications because the GNR non-identification result doesn't necessarily apply when production function parameters are restricted (e.g. constant) over time, and 2) it means that

[32] In some situations it might make sense to use both instruments. For example, suppose one can speculate on each j's set of competitors, but is not sure that these sets are exactly correct. Then the instruments $\boldsymbol{f}_{-jt}$ and $\boldsymbol{\omega}_{-jt-1}$ (which are potentially mismeasured) and the instrument v_{jt-1} (which is inherently inefficient because it depends on $\boldsymbol{f}_{-jt-1}$ instead of $\boldsymbol{f}_{-jt}$) will typically contain distinct information and thus be useful together.

the second stage is linear in the parameters and can be estimated in closed form by 2SLS (alleviating the need to ensure correct numeric optimization for many Monte-Carlo replications).

Wanting our simulated datasets to be somewhat representative of real world data, we use the DGP of Collard-Wexler and De Loecker (2023), which is calibrated to existing datasets, to generate the predetermined data for each duopoly market $(f_{j0}, f_{j1}, \omega_{j0}, \omega_{j1}, f_{k0}, f_{k1}, \omega_{k0}, \omega_{k1})$. But since their setup assumes perfect competition, we deviate when generating variable inputs and quantities. In each market and time period we calculate the equilbrium of a Cournot game to calculate variable inputs $(v_{j0}, v_{j1}, v_{k0}, v_{k1})$ and the implied quantities $(q_{j0}, q_{j1}, q_{k0}, q_{k1})$. The demand curve for the Cournot game has constant elasticity -3 and is calibrated to replicate (on average) the quantities in Collard-Wexler and De Loecker. We also briefly consider a monopoly version of this dataset where $(f_{j0}, f_{j1}, \omega_{j0}, \omega_{j1})$ is generated in the same way, and (v_{j0}, v_{j1}) and (q_{j0}, q_{j1}) are set using the monopoly FOC under the same demand curve.

This model can be estimated by 2SLS, i.e. regressing q_{j1} on f_{j1}, v_{j1}, and ω_{j0}, using as instruments f_{j1}, ω_{j0} (since they are in I_{j0}) plus some version of the competitive instruments from Section 4. Our goal is to assess the precision of estimates provided by the various competitive instruments. A first observation is that in the monopoly dataset, there is no identification. Of course, there are no competitors, so no f_k or ω_k to use as instruments, but one can attempt to use v_{j0} as an instrument. But because there is no variation in the level of competition, this DGP falls within the GNR non-identification framework, i.e. v_{j0} is not a useful instrument and the production function is not identified.

Moving to the duopoly situation, we want to compare use of the "true" competitive instruments, i.e. f_{k1} and ω_{k0} (or $\theta_f f_{k1} + \rho\omega_{k0}$), to using what can be thought of as an imperfect version of these instruments, i.e. v_{j0}. As described in Section 4.5, this can be thought of as a comparison of a situation where the researcher observes competitors (and hence can use f_{k1} and ω_{k0}) versus a situation where the researcher does not observe competitors (or is hesitant to identify who is and who is not a competitor), and hence is uses v_{j0} as a substitute to those competitive instruments. Based on our prior results, we expect the production function to be identified in either case - the more interesting question is the relative precision of the estimates in the two situations. By comparing these, we can assess the value (in terms of lower variance estimates) of an empirical researcher knowing the set of competitors of a firm versus not knowing them.

The production function coefficients are in fact identified in either case with our Monte-Carlo datasets (we do 10000 replications, each dataset containing 10000 observations with $\theta_f = 0.4$, $\theta_v = 0.6$, $\rho = 0.7$). Since the coefficients are very precisely estimated and there are no evidence small sample biases (not unexpected since each firm only has one competitor), we only report the *standard deviations* of the estimated coefficient

$\widehat{\theta}_v$ across the replications. These are more interesting as they shed direct light on our question regarding precision. The following table reports these standard deviations for various types of instruments

Instrument z	Just z	Cubic in z	z, $z \cdot f_{j1}$, $z \cdot \omega_{j0}$	Full 2nd order	Full 3rd order
$\theta_f f_{k1} + \rho\omega_{k0}$	0.0158	0.0147	0.0139	0.0124	0.0123
v_{j0}	0.0139	0.0138	0.0138	0.0137	0.0135
$\theta_f f_{k1} + \omega_{k1}$	0.0132	0.0120	0.0117	0.0103	0.0103

The first row uses the true competitive instrument $\theta_f f_{k1} + \rho\omega_{k0}$, i.e. assuming it is observed[33]. The second row uses the imperfect instrument v_{j0}, i.e. it can be done if competitors are not observed. The last row uses $\theta_f f_{k1} + \omega_{k1}$ as the instrument - this is valid under the additional assumption that innovations in ω are uncorrelated across firms in a given market (and like row 1 requires competitors to be observed).

The first column of the table naively uses either $\theta_f f_{k1}+\rho\omega_{k0}$, v_{j0}, or $\theta_f f_{k1}+\omega_{k1}$ (in the 3 rows respectively) as single instruments in just identified 2SLS procedures. Perhaps surprisingly, using v_{j0} as an instrument generates *more* precise estimates (std. dev. 0.0139) than does using the true competitive instrument $\theta_f f_{k1} + \rho\omega_{k0}$ (std. dev. 0.0158). This is because, as mentioned in Section 4, these instruments likely do not enter in linear form, e.g. in 2SLS the endogenous variable v_{j1} might be better predicted by a non-linear function of $\theta_f f_{k1} + \rho\omega_{k0}$ or v_{j0}. We examine this in the next 4 columns, which increasingly add additional higher order polynomials of the original oligopoly instruments (as well as interactions with the "exogenous" covariates f_{j1} and ω_{j0}) as additional instruments in the 2SLS estimation procedure. As can be seen in the last column, once one gets to a full set of third order instruments (including interactions with the covariates), the true competitive instruments do show their superiority. That said, the difference is not that large - the standard deviation of the coefficient only decreases from 0.0135 to 0.0123 when one uses the true competitive instruments rather than simply v_{j0} - suggesting that in some cases there may not be much precision gain from actually observing the identity of competitors. Interestingly, there is substantially more increase in precision in the last row when one makes the addtional assumption that innovations in in ω are uncorrelated across firms in a given market (and hence can use $\theta_f f_{k1} + \omega_{k1}$ as an instrument instead of $\theta_f f_{k1} + \rho\omega_{k0}$)

[33]For this table we simply used the true parameters θ_f and ρ to construct the instrument $\theta_f f_{k1} + \rho\omega_{k0}$, but the results are essentially the same if either 1) we just use f_{k1} and ω_{k0} as 2 distinct instruments (i.e. which uses weights of standard overidentified 2SLS), or 2) do a two step procedure where we first do 1), and then construct instruments for the second step using the first step estimates, i.e. $\widehat{\theta}_f f_{k1} + \widehat{\rho}\omega_{k0}$. This might be different if these were more than duopolies and there was more scope for small sample biases due to many instruments.

Returning to the first column of standard deviations in the table, we find it interesting that the lagged labor instrument is actually superior in terms of precision to $\theta_f f_{k1} + \rho\omega_{k0}$ when naive linear first stage specifications are used. This points to a subtle potential advantage of the lagged labor instrument. What is going on here is that, at least in our experiments, a linear v_{j0} is apparently a "better" functional form as an instrument (i.e. more predictive) than is a linear $\theta_f f_{k1} + \rho\omega_{k0}$. This makes sense since in our monte-carlo, the functional form through which $\theta_f f_{k0} + \omega_{k0}$ affects v_{k0} is exactly the same as the functional form through which $\theta_f f_{k1} + \omega_{k1}$ affects v_{k1}, and hence a simple linear v_{k0} is almost by default a "reasonable" functional form for predicting v_{k1}. This is not necessarily true when using a linear $\theta_f f_{k1} + \rho\omega_{k0}$ as an instrument. Of course, in real world markets, more things may be changing over time than in our Monte-Carlo setup, but to the extent primitives may not be changing much over time, this result may also hold more generally. Of course, in practice one can use higher order terms (with the caveat of potential overidentification and small sample biases), and one may also want to consider using *both* $\theta_f f_{k1} + \rho\omega_{k0}$ and v_{j0} as instruments in the case where competitors are observed. While the latter would add no additional information in our Monte-Carlo experiments (since in our DGP, v_{j0} is only an imperfect version of the optimal instruments, e.g. because it depends on f_{k0} rather than f_{k1}), in practice it might contain additional information - e.g. the previous literature has noted that v_{j0} can also be an informative instrument for v_{j1} if there is serial correlation in unobserved firm-specific input prices (though this generally is only consistent with DP and TC approaches), or if there are dynamic implications of firms' choices of v.

6 Extensions

6.1 Multiple Variable Inputs

We have restricted attention so far to production functions with only a single variable input v_{jt} (though there can be additional variable inputs that enter in Leontief fashion, e.g. in as in the TC approach). We now discuss implications of having multiple variable inputs (that are substitutable, i.e. not Leontief). Regarding the OP/LP-style inversion results discussed in Section 3, the quantity setting and Bertrand-Nash games are more complicated (since now firms are simultaneously choosing levels of *multiple* variable inputs). This would significantly complicate the notation, but intuition suggests that the sufficient statistic results regarding competition should still hold, and that given sufficient smoothness/substitution between the variable inputs, that any of variable inputs could be used for the first stage inversion.[34]

Here we focus more on the implications of multiple variable inputs on the use of oligopoly instruments in Section 4, as the implications are more severe. Consider a Cournot model, for example. If firm j's first order condition depends on competitors only through the scalar Q_{-jt}, then it is hard to see how exogenous oligopoly instruments could move around *multiple* variable inputs independently. In fact, we can show

Theorem 8 *Consider the Cournot model of Section 3.3 and assume a two period duopoly and a production function with no fixed input, i.e.* $q_{j1} = q_v(v_{j1}) + \omega_{j1} = q_v(v_{j1}) + g(\omega_{j0}) + \xi_{j1}$. *This model is not non-parametrically identified if* $\dim(v_{j1}) \geq 2$.

Proof. Consider a best case scenario where ω_{j0} and ω_{k0} are observed by the econometrician. These two scalars are the only exogenous variables in the model, yet there is a single dimensional function $g(\cdot)$ plus at least a 2 dimensional ($\dim(v_{j1})$) function function to be identified. Hence the model fails a basic rank condition for non-parametric identification. ■

While this Theorem obviously uses a very stylized model to keep things simple, we believe the non-identification result could be extended quite generally. The basic point is that standard models of competition restrict firm interactions on the demand side (i.e. only through quantities or prices), so it is hard to think about models where changes in the amount of competition a firm faces will generate usable variation in that firm's mix of variable inputs.

We think this is an important limitation of oligopoly instruments to recognize, but also want to emphasize that oligopoly instruments can be combined with other sources of identification to identify such models with multiple variable inputs. Again, with the DP and TC approaches, v_{jt-1} can be an informative instrument for v_{jt} if there are unobserved, firm-specific, serially correlated input price shocks. Hence, use of both v_{jt-1} and oligopoly instruments could generate identification of coefficients on multidimensional v_{jt}.

Another possibility is to combine oligopoly instruments with a first order condition approach (e.g. like GNR). In a Cobb-Douglas production function, profit maximization implies that ratios of coefficients on two variable (and static) inputs equal the ratio of expenditures on those two inputs. Enforcing this condition using measured expenditures, the muitiple variable input coefficients can be written as a function of just a single unknown parameter, and hence a single oligopoly instrument should provide sufficient exogenous variation for identification. The intuition here is that imposition of the FOCs identifies how individual variable inputs contribute to a variable input "bundle", and the oligoply instruments provide exogenous

[34] This also gives the potential for testing, as in Raval (2023).

variation to identify the effect of increasing or decreasing that bundle. Interestingly, unlike GNR, this FOC approach does not require the assumption of perfect competition in output markets, since it only relies on FOCs to obtain ratios of variable input coefficients, and in obtaining this ratio, demand elasticities cancel out.[35]

6.2 Revenue Production Functions

Next we consider an important issue in the literature, the question of how output is measured. The above analysis assumes that the econometrician observes (up to measurement error) quantities across firms that are comparable - e.g. in the quantity setting game this means that a unit of q_{jt} is the same across firms; in Nash-Bertrand Logit or Nested Logit it requires that, up to logit and nested logit errors, a unit of q_{jt} is the same for two firms with the same product characteristics x_{jt}. It is straightforward to see the problems if this is not the case - for example, suppose we have a Cournot model of bread slice production where one firm produces loaves of bread with 10 slices, and another firm produces loaves of bread with 20 slices (equally sized). If firms report q_{jt} in terms of loaves, then units are not comparable across the two firms. and, e.g. the first firm may erroneously appear much more productive than the second just because it puts less slices in each loaf.

In a case like this, it may be preferable to use revenue of firms R_{jt} (or log revenue r_{jt}) as a measure of output. We now present an extension of the Cournot model of Section 3 that applies to this case. It also can be used in cases where only revenues (and not q_{jt}'s) are observed, along with market level price deflators. We call this model the "Quality Equivalent Cournot" (QEC) model, because it is essentially a Cournot model, but where firms produce units of different quality/quantities.

Inverse demand is again given by $P_t = P(Q_t, Z_t)$ in the QEC model, but Q_t is now measured in *standardized units*, i.e.

$$Q_t = \sum_j \kappa_{jt} Q_{jt}$$

where the κ_{jt} are scaling factors that measure the relative quality/quantity of each firm j's product. This implies that from a consumer perspective, $\frac{1}{\kappa_{jt}}$ units of firm j's product is equivalent to $\frac{1}{\kappa_{kt}}$ units of firm k's product. P_t is now interpreted as the price of a standardized unit in market t.

Importantly, we allow that the κ_{jt}'s are unobserved to the econometrician. This essentially means that measured Q_{jt}'s are uninformative - a firm could have a high Q_{jt} but actually be producing very little in a

[35] It does require firms to be price takers in input markets, which is relaxed below.

substantive sense if κ_{jt} is very small. On the other hand, observed revenues of each firm, i.e.

$$R_{jt} = P_t \kappa_{jt} Q_{jt}$$

are informative, e.g. if $R_{jt} = 2R_{kt}$ for firms j and k in the same market t, then firm j is producing twice the amount of "quality adjusted" units as firm k. While the κ_{jt}'s need not be observed, we do assume that the econometrician observes a market-level price deflator $\overline{P}_t$ that measures the relative price of a standardized unit in different markets t. Since demand is Cournot (after taking into account the κ_{jt}s), this deflator is the answer to a very simple question, i.e. "How much does a dollar purchase of the product in market t vs market s?". Then we can construct deflated revenues

$$\widetilde{Q}_{jt} = \frac{R_{jt}}{\overline{P}_t}$$

and these are now comparable across markets, i.e. if $\frac{R_{jt}}{\overline{P}_t} = 2\frac{R_{ks}}{\overline{P}_s}$ for firms j and k in markets t and s (respectively), then firm j is producing twice the amount of quality adjusted units as firm k.

Given this setup, assuming F_{jt}, ω_{jt}, and κ_{jt} that are known to all firms, consider a Cournot-Nash Equilbrium in firms' choices of variable inputs V_{jt}. This is exactly equivalent to the model in Section 3 replacing quantity Q_{jt} with deflated revenues $\widetilde{Q}_{jt}$, and all the prior results concerning Cournot models hold. Importantly, note that if this is the model, data on deflated revenues is preferrable to data on quantities. Again, there is no need to observe κ_{jt}, and κ_{-jt}'s, and these can be correlated with unobservables. In fact, there is no real compelling reason to even try to observe κ_{jt}'s, Q_{jt}'s, or firm specific prices given their own definition of a unit of output (i.e. non-standardized units) because deflated revenues are the "true" measure of what firms in this model are producing.

6.3 Imperfectly Competitive Input Markets

Our extensions of the OP/LP-style approach in Section 3 make use of the function

$$V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, Z_t) \tag{48}$$

or more generally

$$V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, Z_{jt}) \tag{49}$$

While these are flexible w.r.t. some different models of competitive interactions in output markets, they do place restrictions on the nature of variable input markets[36]. For example, they are consistent with perfectly competitive variable input markets in which all firms in a given market t face the same variable input price, or where each firm j faces a different competitive variable input price. Of course, the market-specific price (or the firm-specific prices), would need to be observed and included in Z_t (or Z_{jt}) respectively.[37] (48) or (49) are also consistent with "monopolistically competitive" input markets, where each firm j faces its own upward sloping supply curve of V, as long as any potential differences in those supply curves across markets or firms is observed and included in Z_t or Z_{jt}.

However (48) and (49) are generally not consistent with variable input markets in which firms in the same market t interact, e.g. where one firm's choice of variable input quantity affects the price of the variable input for other firms in market t. We next sketch out how input market interactions can be incorporated in a similar way to how we treated output markets. Consider a simple Cournot-like (homogeneous) variable input market in which all firms j in market t participate. Assume the following inverse supply curve for the variable input

$$P_t^V = P_V\left(V_t, Z_t\right) \tag{50}$$

where $V_t = \sum_j V_{jt}$ is the total amount of the variable input demanded by all firms in market t. In other words, the price all firms pay per unit of the variable input in market t, i.e. P_t^V, depends on the total variable input purchased by all firms in t. Presumably $\frac{\partial P_V}{\partial V} > 0$. Defining j's competitors' total variable input usage as the scalar $V_{-jt} = \sum_{k \neq j} V_{kt}$, it is straightforward to show that in the models of Section 3, the following equation holds in equilibrium across j and t

$$V_{jt} = V(\omega_{jt}, F_{jt}, X_{jt}, Q_{-jt}, V_{-jt}, Z_t) \tag{51}$$

and can be used for an OP/LP-style first stage inversion. In other words, to accommodate this type of strategic interaction in the variable input market, one simply needs to additionally condition on another scalar, the variable input use of j's competitors. (51) holds because under these assumptions, any two firms facing the same V_{-jt} and Z_t must also face the same residual input supply curve.[38]

[36]There is increased recent interest in the nature of input markets, see e.g. Rubens (2023).

[37]If these variable input prices are not observed, one typically needs to assume they are the same across markets t, in which case they are implicitly part of V (though one might alternatively make the assumption that they are the same only within groups of markets indexed by m, and estimate a separate V_m for each of these groups of markets (e.g. z group contains all geographic markets in a given year)). .

[38]Presumably these results could also be extended to some cases where the definition of input markets does not necessarily

While interesting in its own right, the above discussion also provides a potential resolution to the multiple variable input issue decribed in Section 6.1. Suppose there are two variable inputs, V^1 and V^2, each in which firms compete within market, e.g. facing two separate market supply curves similar to (50). Also suppose that these variable inputs interact differentially with F and ω in the production function (ruling out simple Cobb-Douglas). If this is the case, then through the interaction in input markets, variation in competitors' F's and ω's can potentially *differentially* affect firm j's choices of the two variable inputs. For example, suppose that the production function is such that F relatively increases the marginal product of V^1 (vs V^2), and ω relatively increases the marginal product of V^2 (vs V^1). Then a firm facing competitors with relatively high F's will in equilibrium face relatively higher prices of V^1 and choose relatively less of this variable input, while a firm facing competitors with relatively high ω's will in equilibrium face relatively higher prices of V^2 and choose relatively less of that variable input.[39] In other words, competitive instruments now have the ability to relatively move the two variable inputs and identify coefficients on both of them, avoiding the non-identification result in 6.1. Note that this argument is relevant for not only OP/LP-style approaches, but also for the use of competitive instruments in the TC and DP approaches. And, as usual, these latter two approaches do not require assumptions on the nature of competition in the variable input markets such that an equation like (51) can be derived.

7 Conclusion

We have described how the presence of imperfect competition and potential strategic interactions introduces distinct challenges when estimating production functions. We started by highlighting that some existing approaches to production function estimation cannot completely abstract away from the presence of imperfect competition in the product market. We then extended these existing approaches to accommodate additional oligopoly models commonly used in empirical work by using a sufficient statistic approach, and showed that the presence of such strategic interactions has important benefits in that they introduces additional exogenous variation that can help identify production functions. We also considered how to optimally leverage this exogenous variation, both with and without direct data on firms' competitors, and in Monte-Carlo experiments 1) verified that the existence of strategic interactions can identify production functions that would not otherwise be identified, and 2) assessed the extent to which what applied researchers ob-

correspond to the definition of output markets.

[39]If there are multiple fixed inputs in F, the same argument could be be made if V^1 and V^2 differentially interact with the elements of F.

serve about competition can affect the precision of estimates based on this variation. As usual, it would be interesting to investigate how the approaches here extend and combine with other recent developments.[40]

[40] For example, the models in this paper for the most part assume Hicksian neutral technological change, so it would be interesting to assess them w.r.t. the recent literature on non-Hicksian neutral technological change, e.g. Doraszelski and Jaumandreu (2018), Raval (2019), Zhang (2019), Kasahara, Schrimpf, and Suzuki (2020), Demirer (2020), Raval (2020), Oberfield and Raval (2021), Balat, Brambilla, and Sasaki (2022), Pan (2022), Ackerberg, Hahn, and Pan (2023).

References

[1] Ackerberg, D.; Benkard, L.; Berry, S. and Pakes, A., 2007, 'Econometric Tools for Analyzing Market Outcomes,' in Heckman, J.J. and Leamer, E. (eds.), *Handbook of Econometrics Volume 6A* (North-Holland, Amsterdam, Netherlands).

[2] Ackerberg, D.; Caves, K. and Frazer, G., 2015, 'Identification Properties of Recent Production Function Estimators,' *Econometrica*, 83, pp. 2411-2451.

[3] Ackerberg, D; Hahn J. and Pan, Q., 2022, 'Nonparametric Identification Using Timing and Information Set Assumptions: An Application to Capital-Biased Technological Change,' mimeo, see *https://drive.google.com/file/d/1yM95D-zyiHs85AgtI46FLqGvMxc2V6nA/view.*

[4] Ackerberg, D. 2023, "Timing Assumptions and Efficiency: Empirical Evidence in a Production Function Context", *Journal of Industrial Economics*, 71:3, pp. 644-674

[5] Arellano, M. and Bond, S., 1991, 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,' *The Review of Economic Studies,* 58, pp. 277-297.

[6] Asker, J.; Collard-Wexler, A. and De Loecker, J., 2014, 'Dynamic Inputs and Resource (Mis)Allocation,' *Journal of Political Economy*, 122, pp. 1013-1063.

[7] Backus, M., Conlon, C., and Sinkinson, M.,2021, "Common Ownership and Competition in the Ready-to-Eat Cereal Industry", mimeo

[8] Bajari, P., Benkard, L., and Levin, J., 2007, "Estimating Dynamic Models of Imperfect Competition", *Econometrica*, 75:5, pp. 1331-1370.

[9] Balat, J., Brambilla, I., and Sasaki, Y., 2022, "Heterogeneous firms: Skilled-labor productivity and the destination of exports", working paper.

[10] Berry, S. and Haile, P., 2014, "Identification in Differentiated Products Markets Using Market Level Data", *Econometrica*, 82:5, pp. 1749-1797.

[11] Berry, S., Levinsohn, J., and Pakes, A., 1995, "Automobile Prices in Market Equilibrium", *Econometrica*, 63:4, pp. 841-890

[12] Biondi, F. (2022), "Firm productivity and derived factor demand under variable markups", working paper

[13] Blundell, R. and Bond, S., 2000, 'GMM Estimation with Persistent Panel Data: An Application to Production Functions,' *Econometric Reviews*, 19, pp. 321-340.

[14] Bond, S. and Söderbom, M., 2005, 'Adjustment Costs and the Identification of Cobb-Douglas Production Functions,' The Institute for Fiscal Studies, Working Paper Series No. 05/04.

[15] Bond, S., Hashemi, A., Kaplan, G., and Zoch, P., 2021, "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data", *Journal of Monetary Economics*, 121, 1-14

[16] Bound, J., Jaeger, D., and Baker, R., 1995 "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak", *Journal of the American Statistical Association*, 90, pp 443-450.

[17] Bresnahan, T. 1982, "The Oligopoly Solution is Identified", *Economics Letters*, 10, pp. 87-92

[18] Bresnahan, T., 1987, "Competition and Collusion in the American Automobile Industry: The 1955 Price War", *Journal of Industrial Economics*, 35:4, pp. 457-482.

[19] Chamberlain, G., 1987, "Asymptotic efficiency in estimation with conditional moment restrictions", *Journal of Econometrics*, 34:3, pp. 305-334.

[20] Collard-Wexler, A. and De Loecker, J., 2015 "Reallocation and Technology: Evidence from the US Steel Industry", *American Economic Review*, 105:1, pp. 131-171

[21] Collard-Wexler, A. and De Loecker, J., 2021, 'Productivity and Capital Measurement Error,' mimeo, see *https://drive.google.com/file/d/1IEab2YyH1QFSWMa6CYAu1mgFET9rn0dz/view.*

[22] De Loecker, J., 2011, 'Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity,' *Econometrica*, 79, pp. 1407-1451.

[23] De Loecker, J. and Scott, P., 2022, "Markup Estimation using Production and Demand Data. An Application to the US Brewing Industry", Working Paper

[24] De Loecker, J. and Syverson, C., 2021, "An industrial organization perspective on productivity" in *Handbook of Industrial Organization, Vol. 4*, pp. 141-223

[25] De Loecker, J. and Warzynski, F., 2012, 'Markups and Firm-level Export Status,' *American Economic Review*, 102, pp. 2437-2471.

[26] Demirer, M., 2020, "Production function estimation with factor-augmenting technology: An application to markups", Working Paper

[27] Doraszelski, U. and Jaumandreu, J., 2013, ‘R&D and Productivity: Estimating Endogenous Productivity,’ *Review of Economic Studies*, 80, pp. 1338-1383.

[28] Doraszelski, U. and Jaumandreu, J., 2018, “Measuring the bias of technological change”, *Journal of Political Economy*, 126:3, pp. 1027–1084

[29] Doraszelski, U. and Jaumandreu, J., 2023, "Reexamining the De Loecker & Warzynski (2012) method for estimating markups", Working Paper

[30] Dhyne, E., Petrin, A., Smeets, V., and Warsynski, F., 2022, "Theory for Extending Single-Product Production Function Estimation to Multi-Product Settings", Working Paper

[31] Duarte, M., Magnolfi, L., Solvesten, M., and Sullivan, C. (2024) "Testing Firm Conduct", mimeo

[32] Flynn, Z., Gandhi, A. and Traina, J., 2019 "Measuring Markups with Production Data", Working Paper

[33] Gandhi, A., and Houde, J-F, 2021, "Measuring Substitution Patterns in Differentiated-Products Industries", Working Paper

[34] Gandhi, A.; Navarro, S. and Rivers, D., 2020, ‘On the Identification of Gross Output Production Functions’ *Journal of Political Economy*, 128, pp. 2973-3016.

[35] Grieco, P.; Li, S. and Zhang, H., 2016, ‘Production Function Estimation with Unobserved Input Price Variation,’ *International Economic Review*, 57, pp. 665-690.

[36] Hahn, J., Liao, Z. and Ridder, G., 2018, "Nonparametric two-step sieve M estimation and inference", *Econometric Theory*, 34, pp. 1281–1324

[37] Hahn, N., 2023, "Who Is in the Driver’s Seat? Markups, Markdowns, and Profit Sharing in the Car Industry", Working paper

[38] Levinsohn, J. and Petrin, A., 2003, “Estimating Production Functions Using Inputs to Control for Unobservables,’ *Review of Economic Studies,* 70, pp. 317-342.

[39] Kasahara H., Schrimpf, P., and Suzuki, M., 2020 “Identification and estimation of production function with unobserved heterogeneity”. University of British Columbia, mimeo

[40] Matzkin, R., 1994 "“Restrictions of Economic Theory in Nonparametric Methods,” in *Handbook of Econometrics*, Vol. IV, R.F. Engel and D.L. McFadden, eds, Amsterdam: Elsevier, Ch. 42.

[41] Matzkin, R., 2013, "Nonparametric Identification in Structural Economic Models", *Annual Review of Economics*, 5, pp. 457-486

[42] Nagel, R., 1995, "Unraveling in Guessing Games: An Experimental Study", *American Economic Review*, 85:5, pp. 1313-1326

[43] Navarro, S., and Rivers, D., 2018, "Nonparametric Identification and Estimation of the Bias of Technological Change", Working paper

[44] Nocke, V., and Schutz, N., 2018, "Multiproduct-Firm Oligopoly: An Aggregative Games Approach", *Econometrica*, 86, pp. 523-557.

[45] Oberfield, M. and Raval, D., 2021, "Micro data and macro technology", *Econometrica*, 89:2, pp. 703–732.

[46] Olley, S. and Pakes, A., 1996, "The Dynamics of Productivity in the Telecommunications Equipment Industry,' *Econometrica*, 64, pp. 1263-1295.

[47] Pakes, A. and Olley, S., 1995, 'A Limit Theorem for a Smooth Class of Semiparametric Estimators,' *Journal of Econometrics*, 65, pp. 295-332.

[48] Pan, Q., 2022, "Identification of Gross Output Production Functions With Nonseparable Productivity Shocks", Working paper

[49] Petropoulous, W., 2000, "Industry Productivity Dynamics with Capacity Utilization and Differentiated Products", University of Michigan Ph.D

[50] Raval, D., 2019, "The micro elasticity of substitution and non-neutral technology", *The RAND Journal of Economics*, 50:1, pp. 147–167

[51] Raval, D., 2023, "Testing the production approach to markup estimation", forthcoming, *Review of Economic Studies*

[52] Reynaert, M. and Verboven. F., 2014, "Improving the performance of random coefficients demand models: The role of optimal instruments", *Journal of Econometrics*, 179:1, pp. 83-98

[53] Reubens, M., 2023, "Market Structure, Oligopsony Power, and Productivity", *American Economic Review*, 113:9, pp. 2382-2410

[54] Shenoy, A., 2021, 'Estimating the Production Function under Input Market Frictions,' *Review of Economics and Statistics*, 103, pp. 666–679.

[55] Van Biesebroeck, J., 2007, 'Robustness of Productivity Estimates,' *The Journal of Industrial Economics*, 55, pp. 529-569.

[56] Wooldridge, J., 2009, 'On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables,' *Economics Letters*, 104, pp. 112-114.

[57] Zhang, H., 2019, "Non-neutral technology, firm heterogeneity, and labor demand", *Journal of Development Economics*, 140, pp. 145–168

8 Appendix

8.1 Proof of Theorem 4

By construction the system of equations hold for the true $\epsilon_{1t},...,\epsilon_{Jt}$ at the true (observed) vectors q_t, v_t, and f_t To prove that the the true $\epsilon_{1t},...,\epsilon_{Jt}$ are the only solution, suppose there is a second solution $\widetilde{\epsilon}_{1t},...,\widetilde{\epsilon}_{Jt}$ (where at least one $\widetilde{\epsilon}_{jt} \neq \epsilon_{jt}$) . This means that $q_{jt} = v^{-1}(v_{jt}, f_{jt}, \sum_{k\neq j}\exp(q_{kt} - \widetilde{\epsilon}_{kt}), z_t) + \widetilde{\epsilon}_{jt}$ for all j. Switching back to the originally defined v^{-1} (i.e. before subtracting $q(v_{jt}, f_{jt}; \theta)$), this is equivalent to

$$q_{jt} = q(v_{jt}, f_{jt}; \theta) + v^{-1}(v_{jt}, f_{jt}, \sum_{k\neq j}\exp(q_{kt} - \widetilde{\epsilon}_{kt}).z_t) + \widetilde{\epsilon}_{jt} \qquad \forall j$$

This then implies that in addition to the true actual quantities $q^*_{jt} = q_{jt} - \epsilon_{jt}$, there are alternative "true" (not mismeasured) quantities $\widetilde{q}^*_{jt} = q_{jt} - \widetilde{\epsilon}_{jt}$ such that

$$\widetilde{q}^*_{jt} = q(v_{jt}, f_{jt}; \theta) + v^{-1}(v_{jt}, f_{jt}, \sum_{k\neq j}\exp(\widetilde{q}^*_{jt}).z_t) \qquad \forall j$$

and where $\widetilde{q}^*_{jt} \neq q^*_{jt}$ for at least one j. Define $\widetilde{\omega}_{jt} = v^{-1}(v_{jt}, f_{jt}, \sum_{k\neq j}\exp(\widetilde{q}^*_{jt}).z_t)$ for all j - note that at least one $\widetilde{\omega}_{jt} \neq \omega_{jt}$ (since otherwise $\widetilde{q}^*_{jt} = q^*_{jt}\ \forall j$ and $\widetilde{\epsilon}_{jt} = \epsilon_{jt}\ \forall j$). By definition of v^{-1}, a firm with $(f_{jt}, \widetilde{\omega}_{jt})$, facing competitors producing a total quantity of $\sum_{k\neq j}\exp(\widetilde{q}^*_{jt})$, must optimal choose v_{jt} (this relies on the assumptions of the prior section that there is a reaction function (vs reaction correspondence) in the appropriate game). Since this is true for all j, this means that given the vectors of observed fixed inputs f_t and "alternative" productivity shocks $\widetilde{\omega}_t$, the vector of observed variable input choices v_t constitue an equilibrium. This contradicts Assumption (5) that there is a unique vector ω_t that is consistent with the observed vector of variable inputs v_t.

8.2 Proof of Theorem 7

With CD production and constant returns to scale in the variable input, the production function for firm k is

$$Q_{kt} = V_{kt}\exp(\theta_0 + \theta_f f_{kt} + \omega_{kt})$$

Therefore, marginal cost is

$$MC_{kt} = \frac{P_V(Z_t)}{\exp(\theta_0 + \theta_f f_{kt} + \omega_{kt})}$$

Under ETETEILT, firm j believes that firm k will maximize profits assuming a symmetric Cournot-Nash equilibrium where everyone else has marginal cost MC_{kt}. With linear aggregate inverse demand $P_t = \alpha - \beta Q_t$, this means firm j believes firm k's thinks its revenues are

$$R_{kt} = (\alpha - \beta (Q_{kt} + Q_{-kt})) Q_{kt}$$

and marginal revenue is

$$\begin{aligned} MR_{kt} &= (\alpha - \beta (Q_{kt} + Q_{-kt})) - \beta Q_{kt} \\ &= \alpha - 2\beta Q_{kt} - \beta Q_{-kt} \end{aligned}$$

Equating to marginal cost gives the FOC

$$\alpha - 2\beta Q_{kt} - \beta Q_{-kt} = MC_{kt}$$

and solving for the symmetric equilibria

$$\alpha - (J+1)\beta Q_{kt} = MC_{kt}$$

or

$$Q_{kt} = \frac{\alpha - MC_{kt}}{(J+1)\beta} \tag{52}$$

Since firm j believes that each opposing firm k will produce according to (52), firm j's percieved FOC is

$$\alpha - 2\beta Q_{jt} - \beta Q_{-jt} = MC_{jt}$$

or

$$\alpha - 2\beta Q_{jt} - \beta \sum_{k \neq j} \left(\frac{\alpha - MC_{kt}}{(J+1)\beta} \right) = MC_{jt}$$

Solving out for firm j's optimal Q_{jt} gives

$$
\begin{aligned}
Q_{jt} &= \frac{\alpha - \beta \sum_{k \neq j} \left(\frac{\alpha - MC_{kt}}{(J+1)\beta} \right) - MC_{jt}}{2\beta} \\
&= \frac{\alpha - \sum_{k \neq j} \left(\frac{\alpha - MC_{kt}}{(J+1)} \right) - MC_{jt}}{2\beta} \\
&= \frac{\alpha - \frac{J-1}{J+1}\alpha - \frac{1}{J+1} \sum_{k \neq j} MC_{kt} - MC_{jt}}{2\beta} \\
&= \frac{2\frac{\alpha}{J+1} - \frac{1}{J+1} \sum_{k \neq j} MC_{kt} - MC_{jt}}{2\beta} \\
&= \frac{2\frac{\alpha}{(J-1)+2} - P_V(Z_t) \frac{1}{(J-1)+2} \sum_{k \neq j} \frac{1}{\exp(\theta_0 + \theta_f f_{kt} + \omega_{kt})} - MC_{jt}}{2\beta}
\end{aligned}
$$

which depends on $\boldsymbol{\omega}_{-jt}$ and $\boldsymbol{f}_{-jt}$ only through their dimension $J-1$ and the scalar $\sum_{k \neq j} \frac{1}{\exp(\theta_0 + \theta_f f_{kt} + \omega_{kt})}$. As noted in a footnote in the main text, if instead of being constant across markets, one assumes the inverse demand intercept α scales in $J+1$ across markets, i.e. $\alpha = (J+1)\widetilde{\alpha}$ (which would be consistent with more firms entering stronger markets), then the last line of the above reduces to

$$
Q_{jt} = \frac{2\widetilde{\alpha} - P_V(Z_t) \frac{1}{(J-1)+2} \sum_{k \neq j} \frac{1}{\exp(\theta_0 + \theta_f f_{kt} + \omega_{kt})} - MC_{jt}}{2\beta}
$$

which depends on $\boldsymbol{\omega}_{-jt}$ and $\boldsymbol{f}_{-jt}$ only through the scalar $\frac{1}{J+1} \sum_{k \neq j} \frac{1}{\exp(\theta_0 + \theta_f f_{kt} + \omega_{kt})}$ (i.e. regardless of the number of competitors).