Handbook of Production Economics, Subhash C Ray, Robert G Chambers, Subal C Kumbhakar (eds.), New York: Springer, I (forthcoming, 2019) : 1-48.

Market Structures in Production Economics

Devin Garcia

Ernst and Young, LLP-Houston

Levent Kutlu

Georgia Institute of Technology

Robin C. Sickles

Rice University

This Version

April 18, 2018

Handbook of Production Economics, Volume 1 **(Springer)**

Subhash C. Ray, Robert G. Chambers, and Subal C. Kumbhakar (Editors)

1. Introduction

One of the important purposes of industrial organization is to understand the environments i.e., the market structures, in which the firms interact with each other, their customers, and potential entrants as well as the implications of these environments for market outcomes. Market structure is a function of the number of firms in market, the concentration of the market, the technology, the nature of products, and the presence of information asymmetry between firms and customers. These factors help us to describe the market structure and examine how it relates to firm conducts, firm efficiency, and market performance. Issues that are taken up by this literature involve whether collusion and higher prices are facilitated in more concentrated markets and whether market concentration affects firms' research and development (R&D) or other types of investments. Among other benefits, an understanding of such relationships guides the policy makers to determine whether regulatory actions are needed and, if needed, how particular regulations may be implemented. Often, economic policies are implemented at a level that affects a wide range of industries, which requires economists to develop tools that apply in a broad range of settings and industries. However, it is also essential to understand industry-specific differences and industry or even firm-specific predictions as in merger cases. Therefore, not surprisingly, economists have developed tools that can be useful in both these settings. The *structure-conduct-performance* (SCP) paradigm is an early descriptive literature that provided many of the stylized facts about market behaviors. The *bounds approach* concentrates on making predictions that can hold across a broad range of industries. This is achieved by aiming conclusions based on minimal assumptions. On the other hand, *structural methods* rely on game theoretic modeling of structure and generally focus on specific industries. Since the focus is a specific industry, stronger assumptions that are compatible with the industry-specific properties are acceptable.

In our Handbook Chapter, we begin with a discussion of the SCP paradigm and related literature, which is followed by a description and literature review on the bounds approach. Then, we briefly talk about commonly used fundamental market structures. In this section, we illustrate how different combinations of various standard concepts are combined to describe market structures. At the end of this section, we briefly discuss additional potential weaknesses of the SCP approach. We follow this section with a description and brief literature review of the conduct parameter approach, which is mainly used in the market power literature. Dynamic environments and markets with product differentiation play important roles in defining market structures. Hence, we provide additional information about models with dynamic environments and models with product differentiation in separate sections. We finalize our review by discussing in depth the literature on market structure on innovation, borrowing from perspectives of Schumpeter. As innovation is a most important issue in the ongoing public policy debates about productivity slowdowns and the related problem of income inequality, this seems particularly relevant to a Handbook Chapter that addresses market structure in a *Handbook of Production Economics*. Our last section concludes.

2. The Structure-Conduct-Performance Paradigm

The structure-conduct-performance paradigm proposes a one-way causation path that starts with market structure's effect on firm conduct, followed by the effect of firm conduct on market performance. As mentioned in the introduction, market structure is a function of number of factors, such as the number of firms in a market, the concentration of a market, technology, the nature of products, and the presence of information asymmetry between firms and customers. A general approach taken by SCP studies is describing market structure through market concentration. It is argued by SCP that in highly concentrated markets the firm's conduct is likely to be more collusive, which would lead to higher prices and thus higher profits. A potential reason for high concentration is entry barriers. One particular entry barrier is the presence of scale economies. Bain (1956) argues that the need of a firm to be large in order to obtain productive efficiency creates an entry barrier. Bain (1956) showed that there is a significant correlation between scale economies and market concentration. This result is consistent with the idea that scale economies facilitate entry barriers. The follow up studies by Bain and others showed that scale economies is not always the only driving force for high market concentration. For example, in markets with intensive investment in advertisement and/or R&D, it is possible to observe high market concentrations.

In the SCP literature, the hypothesis that increased concentration leads to higher prices is strongly supported by both empirical and theoretical studies. Hence, it is argued that concentration is bad for consumers, which historically paved the way for further antitrust legislation. However, a similar strong relation could not be deduced for market concentration and profitability. In particular, the empirical results for the relationship between market concentration and profitability are mixed. One difficulty in examining this relationship stems from measurement and interpretation issues (Fisher and McGowan, 1983). Another difficulty is that cross-sectional data used in these studies frequently come from different industries. This is problematic due to the fact that demand and supply conditions in these industries may be substantially different. The final criticism for SCP studies is that market structure is assumed to be exogenous. Demsetz (1973) argues that positive correlation between market concentration and profits may be due to efficiency differences of the firms. That is, those firms with superior efficiencies can gain market share and this may lead to higher market concentration levels. Weiss (1989) argues that the studies that focused on concentration and price do not face all the criticisms that market concentration and profit studies receive. For example, many of the price-concentration studies use data across specific local markets within an industry. However, as pointed out by both Bresnahan (1989) and Schmalensee (1989), many of the price-concentration studies suffer from serious endogeneity issues. More precisely, the unobserved demand and supply shocks in a market may not only affect prices but also the market structure (e.g., entry barriers may be affected through changes in costs).

3. The Bounds Approach

A summary of the bounds approach along with technical details and some empirical examples is provided by Sutton (2007). Ellickson (2015) is another study that provides a summary of this approach, which is presented in a somewhat less technical way. The bounds approach is developed in Sutton (1991, 1998). The underlying idea of bounds approach is identifying strong mechanisms that can characterize the market outcomes across a broad range of environments. Hence, the approach aims to make as few assumptions as possible about market structure to generate general testable predictions. Ideally, these assumptions should be based on strongly confirmed empirical regulations to reach a reasonable conclusion. This approach deviates from the notion of a fully specified model, which leads to an (unique) equilibrium outcome. Different models that have been proposed in the literature lead to different equilibrium outcomes, and the bounds approach aims to find the bounds for these outcomes. Therefore, rather than providing a single equilibrium outcome, the approach provides a set of equilibrium outcomes that are reasonable.

The simplest version of this approach involves two stages. In the first stage, the strategic and forward-looking firms decide about entry/investment, and in the second stage they compete on prices. One of the key differences in this setting from SCP is that entry is endogenous. Two fundamental assumptions for the equilibria are that they must satisfy *stability* and *viability*. Stability assumes that any firm that chooses to not enter expects negative profits if they enter the market. Viability assumes that any firm that chooses to enter the market expects non-negative profits if they enter the market.

In general, the bounds approach leads to different conclusions depending on whether advertising and R&D plays an important role in the market. Sutton shows that when advertising and R&D are relatively less important for an industry, the market size negatively affects the minimum level of concentration. In other words, the lower bound for concentration decreases as the market grows larger. This result, however, does not imply a functional relationship between concentration and market size. That is, the result does not rule out the possibility of a positive relationship between market size and concentration. What Sutton finds is a relationship between the lower bound of concentration and market size. The vagueness of this result is an artifact of general assumptions applied.

Another related question that is answered by Sutton is the relationship between the strength of competition and the lower bound for concentration. He finds that, for a given market size, an intensification of price competition will shift the lower bound of concentration upwards. Sutton (2007) argues that there are two candidate mechanisms that may lead to this increase in concentration. Some firms may exit and/or consolidate via mergers and acquisitions. An important policy implication of this finding is that regulations which aim to increase competition may indeed lead to market structures that are more concentrated. Of course, as explained earlier, the bounds approach does not provide functional relationships and thus this result does not necessarily imply that an increase in competition would lead to more concentrated market structures.

For industries where advertising and R&D are relatively more essential, Sutton finds that the share of the largest firm remains positive as the market size grows. Hence, the concentration is bounded below. This result contrasts with the findings for industries where advertising and R&D are relatively less crucial. In these industries, larger markets are still more profitable than small ones, but also a larger market size increases the incentives to increase the amount of fixed investments. This in turn prevents market concentration from falling indefinitely. These results explain why in some industries firms have incentives that result in a growing amount of unprofitable sunk investments.

One interesting yet difficult question is whether the results of standard multi-stage models are carried out to the dynamic games framework. Sutton (1998) (Chapter 13) presents a dynamic game in this context. Sutton (2007) argues that the multi-stage game framework excludes certain kinds of equilibria that may arise in a dynamic framework, which is not surprising given the richer essence of dynamic games. Sutton explains that this leads to the appearance of what he refers to as an "underinvestment equilibria." In a dynamic setting, firms may underspend on R&D. This is because of strategies that result in a firm increasing R&D spending at time *t+1* after their rival has increased its R&D spending at time *t*. Nocke (2007) shows that this kind of equilibrium may happen in dynamic games in which the firms can react to rivals arbitrarily quickly. As a consequence, the lower bound to concentration reduces relative to the multi-stage counterpart of a dynamic game.

Symeonidis (2000, 2001) tests some of the important predictions of Sutton via systematic tests that take advantage of a natural experiment, namely a change in British competition law that took place in the 1960s. It turns out that as the laws strengthened, market concentrations generally increased in manufacturing industries. Ellickson (2007) provides empirical support from the U.S. supermarket industry. The U.S. grocery industry is not very concentrated at the national level and even less so at the city level (generally dominated by 3 or 4 chain firms). Ellickson (2007) argues that while R&D and advertising play relatively small roles, supermarkets invest competitively in information technology to increase product variety. As a result of these investments, the number of local firms is limited.

Marin and Siotis (2007) examine Sutton's results in the chemicals industry. The advantage of this study is that the chemicals industry is comprised of many distinct markets in R&D intensity. Moreover, except for pharmaceuticals, advertising expenditures are low. Hence, this allowed them to concentrate on R&D aspects. As in many market structure studies, this study encountered the difficulty of defining the markets and products. Their results support the findings in Sutton (1998) that the lower bound on market concentration is higher and increasing in product concentration where markets have relatively high R&D intensities.

We finalize this section by pointing out that although the bounds approach of Sutton is based on minimal assumptions, the key predictions that are described above depend critically on some of these assumptions. In particular, the assumption that the game form is exogenous to the model may be challenged. That is, the assumption that the form of the game remains unchanged as the market size changes may not be plausible in every context.

4. Commonly Used Basic Market Structures

The structure-conduct-performance paradigm generally uses market concentration measures to describe market structure. However, some of the commonly used economic models provide more precise descriptions about market structures. In this section, we present a very brief summary of these fundamental market structures that constitute the core of more detailed and potentially more technical market structure representations.

The market environment may depend on the nature of products (e.g., homogenous vs differentiated products), the choice of strategy variables (e.g., quantity vs price), timing of the strategic decisions, presence of conditions that allow price discrimination or not, etc. The simplest market structure models are given in the context of homogenous goods, single price, and static environments. One widely used market structure type is *perfect competition*, which is used as a benchmark model in many settings. In this market structure type, a firm is said to be competitive if it believes that it cannot affect the market price. The importance of this market structure type is that the first-best welfare outcome is achieved. Under this framework firms do not find a room for strategic interactions. Therefore, the firm’s residual demand curve is flat.

At the other extreme, another benchmark market structure type is *monopoly* where there is only one producer/seller for a product(s). It is possible to add some more structure to this market structure type by providing more details about the costs and/or technology, whether the monopolist can and is willing to price discriminate by some specific price discrimination mechanism or not, the nature of the product(s), etc. The base version of monopoly sets one price and sells a perishable good. If average costs are declining over all meaningful quantity ranges, the most efficient outcome would be a single firm to produce all output (e.g., public utilities).

This version of monopoly is said to be a *natural monopoly*. In general, this happens in markets where the fixed costs are very high and marginal cost is relatively low.

In some market environments, the monopoly may find ways to implement some form of price discrimination. The essential ingredient for price discrimination is the ability to identify the customer's types (e.g., their valuations) and ability to prevent arbitrage opportunities. Broadly, there are three types of price discrimination. *First degree price discrimination* (personalized pricing) happens when the firm has perfect knowledge of the valuations of the customers. In this framework, the firm charges different prices for each customer based on their valuations. Total welfare in this setting is the same as in perfect competition. The distinction is that all the welfare goes to the monopoly. While initially this type of price discrimination seemed to be an abstract concept that is hard to find in real world scenarios, recent advancements in machine learning algorithms and big data make this type of price discrimination in some market structures closer to reality than fiction. With *second degree price discrimination* (menu pricing), the firm creates slightly different products (e.g., different amounts of the same product) for the purpose of price differentiation. Finally, *third degree price discrimination* (group pricing) divides the market in segments and charges the same price everyone according to segment. The reason for such segmentation is that the firm may not have the tools to identify individual valuations of the customers. Hence, the firm rather identifies groups in an attempt to get a heuristic approximation that simplifies the identification problem. Two obvious examples of *third degree price discrimination* are student discounts and geographical price discrimination. Armstrong (2006) and Stole (2007) survey this literature and provide details of advancements in the price discrimination literature.

If a product is a durable good rather than perishable good, this would have serious implications for the monopolist. One important implication of market structures with durable goods is that the monopolist needs to take the future into account. In a typical market with perishable goods, it is possible to use static models; however, this is inappropriate with durable goods. Coase's conjecture claims that if consumers do not discount time heavily and expect the price to fall in the future then this would affect the current demand negatively. This in turn forces the monopoly to charge a lower price compared to what it would charge for a perishable good. Some examples that may solve consumers' lower future price expectation problem include introducing capacity constraints, announcing future prices, or renting. Whether this conjecture holds or not is discussed by many researchers and has led to a large literature (Fudenberg and Tirole, 1991).

So far the market structure examples that we provided were either perfect competition or monopoly. More interesting cases can be explored in market structures with imperfect competition. One important distinction of monopoly and imperfect competition models is that for the latter, in general, the choice variable plays a more important role. That is, in imperfect competition models the outcomes may change substantially depending on whether the firms choose quantity or price. Similarly, the timing of a firm's actions plays an important role, e.g., simultaneous vs sequential.

We first start with imperfect competition market structures where the firms choose quantities. *Cournot competition* assumes that there are multiple firms in a homogenous product market that choose quantities to maximize profit. The equilibrium output is determined such that no firm can increase its profits by changing its output level if other firms produce Cournot output levels. That is, firms maximize their profits in accordance with a Nash equilibrium. A characteristic of this market structure is that an increase in another firm's output leads the firm produce less. In other words, the best response functions are downwards sloping. When the number of firms is 1, this market structure coincides with the (based) monopoly structure. When the number of firms goes to infinity, the outcome of this market structure converges to that of perfect competition.

In the Cournot competition model firms act simultaneously. If one firm has some advantage to enable it to move first, then this would change the market outcomes and lead to another market structure known as Stackelberg competition. The simplest version of this model assumes that there are two firms in the market that compete on quantity: the leader and the follower. The leader moves first and after observing the leaders action the follower moves. It is assumed that the leader knows *ex-ante* that the followers observe their actions. Moreover, the follower must not have the ability to commit to a future non-Stackelberg follower action, and this has to be known by the leader. One reason why a leader might have a first-mover advantage is that it may be the incumbent monopoly in the industry, and the follower is a potential new entrant. Hence, the Stackelberg competition model and its variations play an important role in describing markets that face potential entrants.[1]

Similar to monopoly setting, it is possible to add more details about the market structure to the quantity competition based models. For instance, Stole (2006) gives an example of third degree price discrimination in a quantity choice setting. Similarly, Hazledine (2006, 2010) and Kutlu (2012) incorporate a form of second degree price discrimination to the Cournot competition framework. They show that, under linear and some non-linear demand functions, the quantity weighted average price does not depend on the extent of price discrimination. Kutlu (2009) incorporates the same price discrimination type to the Stackelberg competition framework where the demand is linear and costs are symmetric. He shows that the leader does not use price discrimination but the follower does. The leader directs all of its first mover advantage to attract the highest value consumers.

Whether firms actually set quantities or not depends on the particularities of relevant markets. *Bertrand competition* assumes that the firms set prices rather than quantities. In a homogenous goods market with Bertrand competition, the consumer is assumed to buy from the firm with the lowest price. When the firms charge the same price, we need a sharing rule. A sensible sharing rule is equally dividing the demand among those firms that charge the smallest price. Under these conditions, if the marginal costs are the same for the firms, the market price reduces to the marginal cost. This result is paradoxical because even with two firms the market price becomes the competitive price. Based on the empirical evidence in the markets with small number of firms the price-cost markup is positive in general. Hence, in the literature this result is referred to as the Bertrand Paradox. Hence, when firms choose prices, the market outcome is materially different compared to Cournot competition.

When capacity constraints are introduced to the price choice framework, it is possible to achieve outcomes like Cournot competition (Kreps and Scheinkman, 1983). In particular, Kreps and Scheinkman (1983) show that it is possible to get Cournot-like outcomes if the firms first choose capacity and then prices. This suggests a potential solution to the Bertrand Paradox. However, their result critically depends on the rationing rule that is used. A technical aspect of

[1] In other contexts, a seminal work on static entry models is Bresnahan and Reiss (1990), which examines the strategic entry decisions of small retail firms.

Bertrand competition is that the equilibrium may not exist if the marginal costs are not constant. This leads to alternative models such as the supply function equilibrium of Klemperer and Meyer (1989). One advantage of the supply function equilibrium approach compared to Bertrand and Cournot games is that in this setting a firm adjusts to the uncertainty in an optimal way given the other firms' behavior.

Until now, we have concentrated on market structures with homogenous goods. However, there are many markets where such an assumption may not be realistic. That is, with rare exceptions, the goods vary in features, quality, location, etc. Also, although Kreps and Scheinkman (1983) provide a plausible solution to the Bertrand Paradox, a more applicable solution is to introduce differentiated products. Product differentiation allows the consumers to choose the product variations close to their tastes. The greater the variety of the products sold in the market, the more likely consumers would find a better match to their ideal preferences. Hence, product differentiation helps the firm to increase the potential amount of goods that they can sell. However, product differentiation comes with the potential cost that firm(s) may have to worry about determining the optimal degree of product differentiation.

There are two ways to achieve product differentiation. In *vertical differentiation,* all consumers have the same preference ranking of the products given that they are charged the same price. In *horizontal differentiation,* consumers do not have the same preference rankings of the products even when the products are charged the same price. In the vertical differentiation setting, consumers may rank the products, say, based on quality. With horizontal differentiation, differences of the products may be due to factors other than quality such as the color, etc. In the literature, markets with vertical differentiation are generally examined by models that account for quality differences of the product variations (e.g., Gabszewicz and Thisse, 1979, 1980 and Shaked and Sutton, 1982, 1983). A common way to model horizontal differentiation is using the variations of linear city model of Hotelling (1929) and circular city model of Salop (1979). In these models, the distinct preferences of consumers for the products are modelled through a parameter called transportation cost. If a variation of the product is not close to a consumer's taste, then the transportation cost of this individual would be high for this product variation.

Many of the current empirical studies choose to model horizontal product differentiation via discrete choice models. In these models, it is important that consumers have heterogeneous preferences and choose only one product (out of a set of finitely many products) to ensure a smooth aggregate demand curve, which seems to be a requirement that is naturally satisfied in most cases. Later, we will provide a more detailed discussion for these types of models.

As we discussed in the SCP section, market structures may be described by market concentration measures such as four-firm concentration measures or Herfhindahl-Hirschman Index (HHI). For example, the Department of Justice Merger Guidelines uses HHI in their merger evaluation analysis. One problem using concentration measures is that it may be difficult to come up with a proper market definition where the market shares are calculated. Another difficulty is that market shares are highly imperfect at describing market dynamics. Thus, they do not give us a complete or precise description of the market structure. For example, two homogenous product industries with the same market concentration measures may have totally different characteristics. In particular, in one industry the market structure may be described with Cournot competition and in the other industry it may be Bertrand competition. Similarly, in one industry the market structure may be described by uniform-pricing firms, and in the other industry the market structure may be described by price discriminating firms. Hence, using HHI as a measure of competition in a cross-industry study may not always be the best way to proceed. In such cases, the models that describe the market by more precise structural information can be used as alternative options.

In the next section, we consider the conduct parameter (or conjectural variations) approach that enables researchers to model market structures by “generalizing” some of the equilibrium concepts that are used in the commonly used models described above.

5. Conduct Parameter Approach

An exhaustive survey of the conduct parameter approach along with technical details and some empirical examples is provided by Perloff et al. (2007). Bresnahan (1989) is another study that provides a detailed summary of this approach. Some of the earlier works that use this approach are Gollop and Roberts (1979), Iwata (1974), Appelbaum (1982), Porter (1983), and Spiller and Favaro (1984).

As mentioned earlier, the focus of the SCP literature is the cross-sectional study of many industries. The conduct parameter approach, on the other hand, concentrates on a single industry in an attempt to estimate a conduct parameter that characterizes firm behavior. Hence, conduct parameter models use economic theory to guide the empirical model specification while concentrating on a single industry. These empirical models rely on the theory of *conjectural variations* to estimate conduct parameters, which is mainly used as a measure of market power.

Based on the conjectural variations interpretation, the conduct parameter measures the market power of firms in a market in a fairly general way by allowing equilibrium outcomes that may not be supported by the standard equilibrium concepts such as Nash equilibrium. For example, in the standard Cournot model, the conjecture is that the firms will have zero reaction; yet conjectural variations theory allows more general types of reactions. Basically, the conjectural variations of the firms determine the slopes of their reaction functions. Hence, similar to the common models that we described, the researchers may add some structure (e.g., capacity constraints, dynamic factors, price discrimination, etc.) to the model that describes the market structure in a market; but at the same time, the researcher may also be agnostic about the firms' competitive behavior, i.e., the firm conduct, and estimate it using the available economic data.

For example, Puller (2009) incorporates capacity constraints; Puller (2007) and Kutlu and Sickles (2012) incorporate dynamic strategic factors; Graddy (1995) and Kutlu and Sickles (2016) incorporate price discrimination in their conduct parameter models. Corts (1999) and Kutlu and Sickles (2016) state that in the language of conjectural variations theory the conduct is described in terms of firms' conjectural variations, which are their expectations about other firms' reactions. Conduct parameter models use the conduct parameters to represent the conjectural variations of the firms. Based on this interpretation, the conduct parameter can take a continuum of values. It is important to note that the conjectures do not refer to what firms believe will happen if they change their quantity levels. Rather, what is being estimated is what firms would do as a result of their expectations. Hence, as argued by Corts (1999), the conduct parameter can be estimated "as if" the firms are playing a conjectural variations game, which would reveal the price-cost margins.

However, some researchers may not be comfortable with the idea of a conjectural variation that allows non-standard equilibrium outcomes, e.g., equilibria other than the Bertrand, Cournot, collusion, etc. Although some other researchers have argued that the folk theorem allows a range of conduct parameter values that are consistent with Nash equilibrium, they could not make a strong case against critics that question the validity of using a static model to represent a dynamic game. Hence, some researchers prefer to view conduct as a parameter that can take values consistent with existing theories, which would be estimated in the conduct parameter model. That is, the estimated conduct parameter value can be used to categorize the competitive behavior of firms by using statistical tests. In many cases, the researcher would face more than two alternative models to pick from. For example, the researcher may need to test whether the market outcome is consistent with perfect competition, Cournot competition, or monopoly. Therefore, non-nested hypothesis tests (e.g., Vuong, 1989) can be used as in Gasmi et al. (1992). Due to its simplicity, some researchers may prefer to choose a compatible model by making pairwise comparisons using the standard statistical tests similar to pairwise model comparisons done in Bresnahan (1987).

Besides the interpretation related criticisms, the standard conduct parameter models that do not incorporate the dynamics specifically are intrinsically static. Corts (1999) argues and illustrates by an example that this may lead to severely mismeasured market power estimates. Puller (2009) and Kutlu and Sickles (2012) present conduct parameter models that are robust to the criticism of Corts (1999). That is, they offer general empirical models that allow the consistent estimation of the parameters of the model (including the conduct parameter) that is robust to efficient collusion. As argued by Puller (2009), while these models are robust to efficient tacit collusion, they may not be robust to other forms of dynamic solutions. Nevertheless, these models nest the static scenario in a testable way, and it would be extremely difficult if not impossible to design economic models that are robust to every, or even most, dynamic market behaviors. More details about dynamic market structures in other contexts are provided in the next section.

One of the distinctions of conduct parameter models from the standard economic models is that the identification requires an extra effort. In particular, when estimating a demand-supply system of equations, the researcher needs to be more careful compared to standard demand-supply models. The source of this problem is that not every functional form choice for demand and marginal cost functions enables separate identification of the marginal cost and conduct parameter. If the functional form choices are not carefully done, it would be possible to confuse competitive markets with high marginal cost and collusive markets with low marginal cost.

Lau (1982) and Bresnahan (1982) provide some conditions for identification in the conduct parameter setting. Bresnahan (1982) suggests that this identification issue can be solved by using more general demand functions so that the exogenous variables do more than parallel shifts, i.e., change the demand slope by rotations. Hence, the rotations around the equilibrium point would identify the conduct parameter. This can be achieved by including an interaction term with the quantity variable. However, Shen and Perloff (2012) show that such rotations may cause some multicollinearity issues.

Another, potentially more realistic, identification approach would be the non-parametric structural identification approach in Brown (1983), Roehrig (1988), and Brown and Matzkin (1998). Recently, Orea and Steinbucks (2018) and Karakaplan and Kutlu (2017) proposed conduct parameter models that can be estimated using stochastic frontier approaches. The advantage of these methods is that they model the firm and time specific conduct parameters as random draws from a doubly truncated normal distribution. Hence, in contrast to existing conduct parameter models, they use skewness of the distribution of conduct parameter in order to identity marginal cost and conduct parameters separately without requiring some of the strong functional form restrictions on the demand and marginal cost functions. In another context, Kumbhakar et al. (2012) also use the stochastic frontier approach to estimate market powers (i.e., markups) of firms. Their approach allows estimation of market power even when the input price data are not available. Moreover, their method can reliably estimate market power with or without constant returns to scale assumption.

6. Dynamic Market Structures

Perloff et al. (2007) distinguish two types of properties that affect the dynamics of market structure of a market: *fundamental* and *strategic*. If the dynamics of the market structure is due to a stock variable that affects future profits, they call this type of reason fundamental. If the dynamic interactions of firms stem from the beliefs that the rivals will respond to current actions, they call this type of reason strategic.

In a dynamic setting, the optimization condition should be modified so that price equals *full marginal cost*, where the full marginal cost equals the sum of marginal cost and a term that is a function of shadow value of the constraints that the firms face. If the reasons are strategic, this can be a function of the shadow value of the incentive compatibility constraints due to cooperation (e.g., Puller, 2009 and Kutlu and Sickles, 2012). If the reasons are fundamental, this can be a function of the shadow value of a stock (e.g., Pindyk, 1985 and Puller, 2007). The stock variable can be amount of natural resources, knowledge, or a quasi-fixed input. A common example of a quasi-fixed input is capital, as in many applications it is more expensive to make quicker adjustments in the capital. This is an example of production-related fundamental reasons (e.g., Reynolds, 1987 and Driskill and McCafferty, 1989).

Any market structure that involves a quasi-fixed input requires that firms solve a dynamic optimization problem. The reason is that the quasi-fixed input affects the current profits and future levels of quasi-fixed input, which in turn affects future profits. Therefore, the optimal level of investment path depends on the current period's quasi-fixed input amount and the firm's belief about future factors such as input prices. Similarly, in markets where advertising is a relatively more important tool that can change demand, the firms face a dynamic optimization problem. This is an example of demand-related fundamental reasons. Perloff et al. (2007) argues that advertising may create a stock effect by increasing the firms' customers today and in the future. If the firm has a small portion of the potential customers, the value of additional advertisement would be large. If this firm invests a large amount in the current time period, it may boost the current demand by capturing potentially a large portion of the market demand. This behavior, however, affects the need for advertisement in the next period because the firm already has a high demand due to large investment on advertisement in the current period; and thus the value of additional advertisement for the next period may be small. The intertemporal connection of advertisement decisions consequently makes the firm's optimization problem a dynamic one.

For market structures where dynamics play an important role, open-loop equilibria and Markov perfect equilibria are among the most commonly used solution concepts. With Markov perfect equilibria, firms know that rivals will respond to a change in the state variable. On the other hand, in open-loop equilibria, the firm will assume that the rivals will not respond to these changes. Therefore, it seems that Markov perfect equilibria more closely reflect what many would call rational firms' behavior. Perloff et al. (2007) provide detailed examples of those models that use these equilibrium concepts.

7. Market Structures with Differentiated Products

In this section, we concentrate more on the approaches that are used when modelling markets with differentiated products and their estimation. The first approach estimates residual demands for close substitutes. Since the degree of market power of a firm depends on the residual demand elasticity, the residual demand approach may be useful in studies that analyze market power. Some earlier examples that use this approach are Bresnahan (1981, 1987) and Spiller and Favaro (1984). The main difficulty of this approach is that it is not clear how one estimates all own- and cross-price elasticities.

The second approach estimates a neoclassical demand system. The general idea is simultaneously estimating a demand system of goods along with marginal cost functions for each good. Hence, for *n* goods this approach estimates a system of 2*n* equations (*n* for demand *and n* for supply). However, when the number of goods is relatively large, estimation of such a system of equations may be impossible due to data requirements. A solution to this problem is estimating a model with a multilevel demand specification (e.g., Hausman et al., 1994, Hausman and Leonard, 2004). This approach imposes some implicit restrictions on the cross-price elasticities. In particular, changes in the prices of one category do not affect the demand for another category. Another solution is estimating a so-called almost ideal demand system (AIDS). A relatively simple version of this method is the linear approximation of AIDS that is proposed by Deaton and Muellbauer, 1980). This approach uses Stone's (1953) geometric approximation to the price index.

One overlooked issue in the literature is the fact that the estimated demand system actually may not be a complete one by ignoring other goods. It is, however, possible to estimate an incomplete demand system in a way that is consistent with utility maximization (e.g., LaFrance, 1990, 2004). The third and relatively more widely used approach estimates a random parameter utility model. This method solves the potential data requirement problems mentioned above. Moreover, Nevo (2000a) argues that the assumptions for consumers to have preferences so that an aggregate consumer exists and has a demand function that satisfies conditions assumed by economic theory are strong, and many times these assumptions are empirically falsifiable. Hence, using an aggregate model may lead to different conclusions compared to a model that explicitly models individual heterogeneity.

One potential solution is to use discrete choice models, which solve the dimensionality problem by projecting the products into a characteristics space (e.g., Berry et al., 1995 and Nevo 2000a, 2000b, 2001). These models allow the researchers to model the market structure at a very detailed micro level. For example, Nevo (2001) estimates brand-level demand for the ready to eat cereal industry and uses the estimates along with the pricing rules to recover price-cost markups without observing the costs. In contrast to the conduct parameters method, which estimates the firm's conduct along with other parameters, Nevo's approach assumes that the firms compete under Nash-Bertrand setting. Hence, the conduct is exogenously given. He instead uses three hypothetical ownership structures to determine the extent of market power: single-product firms; the current structure observed in the data; and a multi-brand monopoly producing all brands. Based on these different ownership scenarios, he calculates the counterfactual price-cost markups that correspond to each of these scenarios. By using crude measures of actual price-cost markups, he can determine which ownership structure fits better to the observed data. Using the same approach, Nevo (2000b) provides an analysis of merger impact prior to its consummation. This provides an important tool for policy-makers for merger analysis.

8. Market Structure and Market Power

There is a strong connection between market structure and market power. At a high level, market structure creates the environment that forms the base of market power. Although they actually measure market concentration, HHI and, to a lesser extent, concentration ratios are commonly used proxies for market power. The appeal of HHI is that it not only gives more weight to larger firms but it also increases as the number of firms decreases. However, HHI does not consider the particularities of markets. For example, although not very likely, it is possible to imagine relatively competitive markets even with two symmetric firms. Application of the HHI in practice is also problematic since it requires that goods be homogenous. Hence, in markets with differentiated products, the usage of HHI requires additional assumptions, e.g., calculating the market shares based on sales rather than actual quantities.

Another common measure of market power is the Lerner index, defined as the ratio of price-marginal cost markup to price. In a standard static setting, when price equals marginal cost, this measure equals zero, which indicates a perfectly competitive market. One issue with the Lerner index is that efficiency improvements could be mistaken as increments in market power. Koetter et al. (2012) propose efficiency adjusted measures of Lerner index to overcome this issue. Moreover, if the market structure involves dynamics (due to either strategic or fundamental reasons), the relevant marginal cost concept for optimality conditions is the full marginal cost. This suggests, that the Lerner index also must use the full marginal cost in its calculation. Kutlu and Sickles (2012) suggest using efficiency adjusted full marginal cost, which is robust to efficiency and dynamics related concerns mentioned above. Kutlu and Wang (2018) provide a variety of conduct parameter based models in this context.

The Lerner index may not always be easy to estimate due to available data issues. More precisely, the calculation of Lerner index requires knowledge of marginal cost, which is not directly observable and thus requires to be estimated. One way to do so is estimating a cost function and calculating the marginal cost from the parameter estimates of this cost function. This, however, requires total cost data, which is not always easy to obtain. The conduct parameter approach enables the estimation of marginal cost implicitly without using the total cost data. Also, this approach provides an alternative measure of market power, i.e., the conduct parameter. At least some variations of conduct parameters are shown to be highly correlated with the Lerner index. In particular, under some assumptions, the conduct parameter equals price elasticity adjusted Lerner index.

Boone (2008a) proposes a new way to measure market power. His starting point is that there are some theoretical examples (e.g., Rosentahl, 1980; Stiglitz, 1989; and Bulow and Klemperer, 1999) where more intense competition leads to higher price-cost margins. Hence, Boone (2008a) aims to develop a competition measure that is theoretically robust yet requires similar data sets with price-cost margin estimation. Boone (2008a) calls his measure *relative profit differences* (RPD). This measure is defined as follows: Let $\pi(n)$ the profit level of a firm with efficiency level n where higher n denotes higher efficiency. Consider the variable $(\pi(n'') - \pi(n))/(\pi(n') - \pi(n))$ for three firms with efficiency levels $n'' > n' > n$. Boone (2008a) argues that an increase in competition raises this variable in models where a rise in competition reallocates output from less efficient to more efficient firms, which covers a broad range of models. Another related study with a very similar measure is Boone (2008b), which proposes a measure called *relative profits*. This time Boone considers the variable $\pi(n'')/\pi(n')$ for two firms with efficiency levels $n'' > n'$. Boone (2008b) argues that an increase in competition reallocates profits from the less efficient firm to the more efficient firm and thus increases the relative profits measure.

9. Market Structure and Innovation-Studies with no Explicit Treatment for Distorted Production Decisions

Joseph Schumpeter (1942) methodically demonstrated innovation's principal role in advancing economic prosperity. A number of innovations have come and gone that strongly support this claim. The steam engine, light bulb, automobile, airplane, radio, and integrated circuit immediately come to mind—each of these innovations have pushed the economic frontier forward, noticeably improving consumer welfare. Understanding what innovation is and how competition policy could affect it is consequently important. The principal goal of this section is to explore the latter, but we first provide a short summary of the origin of the literature, what innovation is and some of the issues surrounding measurement.

Innovation is defined generally as a new or materially improved product, process, service or business method. What drives it is the profit motive. Economic policy, therefore, can only be effective at fostering innovation if it helps to safeguard or promote the incentive to innovate.

Broadly speaking, one may view the profit motive from two perspectives: (1) the objective and expectation of realizing an economic gain solely on the merits of the innovation, with or without existing competition; or (2) the objective to maintain profitability[2] at a minimum or prevent an economic loss when competition is present. The first approach to innovation is purely entrepreneurial in that the innovator's investment decision is independent of its competitive environment. In contrast, the second approach to innovation is a mixture of entrepreneurship and an effort to survive; i.e., the innovator also takes into account its competitive environment to optimize its investment strategy. This section limits discussion around the second approach to innovation.

Whether or not competition spurs innovation has been a question of significant interest in the industrial organization literature. One may attribute this to Schumpeter (1942) who argues that in a capitalist system, perfect competition is both incompatible with economic progress and inferior to large scale enterprise. Taken at face value, his conclusion might indicate that a more competitive industry stifles innovation, but this is not Schumpeter's intention.[3]

Schumpeter's claim that "big business" is superior to atomistic competition in a capitalist system is auxiliary to his main concern–that capitalism is superior to socialism *because* perfectly

[2] Provided profitability is high enough to fully cover economic costs.
[3] For example, Aghion et al. (2005) define the "Schumpeterian effect" as the effect of increased competition lowering post-innovation rents, thereby reducing the incentive to innovate.

competitive markets maximize social welfare. Schumpeter advanced the idea that capitalism is superior not because of price competition or any particular market structure, but rather because free enterprise begets the introduction of new products, processes, services, and methods of production. If one therefore accepts the premise that economic prosperity is the product of innovation, then perfect competition cannot be the reason for capitalism's success as an economic system. Schumpeter demonstrates this point with the following observations.

The first of these observations is that capitalism is a system of constant and disruptive structural change. Accordingly, it is not appropriate to assess capitalism's performance based on an outcome of static conditions. In particular, while the static outcome of perfect competition is socially desirable at any given point in time, it implies that goods and production methods never change. The static efficiency of perfect competition thus comes at the cost of precluding innovation's promise of dynamic and long-run efficiency.

Schumpeter also juxtaposed the assumption of free entry in a perfectly competitive market against an innovator's expected return on investment. That is, even though free entry is a necessary condition for maximizing social welfare, quick and costless entry into a new market is at odds with an innovator's incentive to invest.[4] In particular, if free entry implies costless and nearly instantaneous replication of the innovator's idea, then the innovator's expected return will not exceed its opportunity cost of investment. Even ignoring the assumption of homogeneous goods and static production technologies, it follows that a firm in a perfectly competitive market has no incentive to innovate.

The final observation is that large scale and anticipated market power facilitate investment in innovative ideas. More specifically, the financial position of a profitable, large firm grants it the opportunity to take on riskier innovative activity than a small firm. And anticipated market power, by way of an intellectual property right, promotes innovative activity by allowing the innovator to temporarily charge a price in excess of what the market would bear. The latter observation, while not novel, is further emphasis by Schumpeter that free and prompt entry is problematic for innovation. Schumpeter additionally advocates for intellectual property

[4] The idea that free entry will quickly erode an innovator's supranormal profit assumes that the innovator's competition is effective; that the idea embedded in the innovation is copied well and easily; and that being a first mover does not confer a material advantage. While this is not always the case (see, for example, Boldrin and Levin, 2008), the topic of intellectual property rights is outside the scope of this chapter.

rights as it permits a firm to better plan for the future. On the other hand, Schumpeter's view that large scale is a boon rather than a detriment to innovation is meant to highlight the limitations of a perfectly competitive firm, specifically limitations for growth due to relatively little capital to finance risky innovative activity, and the relative inability to weather the constant flow of outside innovations.

It bears repeating that Schumpeter's criticism of perfect competition as a model of efficiency is not a statement about the relationship between innovation and *ex-ante* market structure, market power or level of competition. Rather, Schumpeter's intention is to demonstrate that the assumptions of perfect competition are wholly in conflict with innovation, and therefore economic progress. The idea that *ex-ante* market power fosters innovation is nevertheless commonly ascribed to Schumpeter. This is borne out by the multitude of studies that have tested some form of Schumpeter's proposition.

Before going into some detail about the literature alluded to above, we briefly touch on the issue of measurement. A common measure of competition is the Lerner index. The ubiquity of this measure is traced to its relationship with the price elasticity of demand, the source of a firm's market power when price is the only choice variable. Indeed, when a single, profit maximizing firm takes its demand as given, its equilibrium Lerner index will be inversely proportional to its price elasticity of demand. Thus, a larger Lerner index is taken to indicate greater market power or less competition.

There are several reasons why this measure could fail to indicate changes in market power. One possibility is if a firm engages in cost-minimizing behavior. In this case, a temporal increase in the Lerner index would reflect nothing more than the competitive process at work. The life cycle of a product will also affect the margin we observe in the data. For many goods, as a product matures, focus shifts from differentiation and quality improvement to cost reduction (Utterback and Abernathy, 1975). That is, commoditization will take effect in the later stages of a product's life cycle, and margins will generally fall as a result. A firm's margin will also increase if it is able to generate a competitive advantage. For example, in a hypothetical industry of two firms selling the same good, successful product innovation by one of the firms will allow the innovator to command a higher price for its product. The average margin in the industry will increase, but if the competitive advantage possessed by the innovator pushes the non-innovator to invest in developing a better product, then rivalry will have been preserved if not strengthened. Thus, an increase or decrease in margins does not fundamentally indicate a fall or rise in competitive intensity, respectively.

The Herfindahl-Hirschman index is another commonly used measure of competition. The main attraction of this measure is its consistency with the intuition that more concentrated industries are less competitive (i.e., industries with few firms and/or firms with substantial market share are viewed as relatively less competitive). Also attractive is the Herfindahl-Hirschman index's theoretical link to market structure. For example, the Herfindahl-Hirschman index corresponds to monopoly when the market is captured entirely by one firm and perfect competition when there are infinitely many firms with equal shares.

The prevalence of the above measures is based on the presumption that changes in market structure or market power will identify changes in competitive intensity. This presumption is questionable; however, if the defining feature of a more competitive market is greater rivalry to deliver superior products and services. In particular, it does not follow that a more profitable or concentrated industry will be less inclined, on average, to compete vigorously (or vice versa), especially when changes in market structure or market power are measurably small. A more competitive market will arise when, on average, a firm faces a greater risk of falling behind (resp., a greater prospect of moving ahead) its competitors if it offers inferior (superior) value to consumers. That is, a more contestable market will result in greater rivalry. Shapiro (2012) calls this the "contestability" principle, and it is adopted by Garcia (2016) to examine the relationship between competition and innovation.

Application of the "contestability" principle is consistent with antitrust policy.[5] The 2010 Horizontal Merger Guidelines (HMG) of United States Department of Justice and the Federal Trade Commission state the following:[6]

"The unifying theme of these Guidelines is that mergers should not be permitted to create, enhance, or entrench market power or to facilitate its exercise... [where by definition] a merger

[5] Anne Bingaman, a former Assistant Attorney General for the Antitrust Division in the U.S. Department of Justice during 1993-1996, said the following about rivalry and innovation: "The fundamental thesis of strong antitrust enforcement is that rivalry, not market power, fosters innovation and efficiency over the long run... Antitrust has an important role in preserving the rivalry that spurs innovation." See Bingaman (1994).

[6] See HMG (2010) for the complete set of guidelines governing antitrust policy in the United States. Italics are our own and used for emphasis.

enhances market power if it is likely to encourage one or more firms to raise price, reduce output, diminish innovation, or otherwise harm customers *as a result of diminished competitive constraints or incentives*."

The Guidelines identify the root of enhanced market power as any market characteristics that would lessen rivalry or the incentive to compete. While it is possible that enhanced market power could manifest as greater profitability or market concentration, the converse is not necessarily true. In contrast, any policy that would serve to make a market more contestable, all else equal, will intensify competition.

The measurement of innovation is equally important. The four defining features of innovation are: investment; expansion of revenue and profit; rare occurrence; and growth in economic output that exceeds input growth. The three most common measures of innovative activity are R&D intensity, patent counts, and productivity growth. We briefly argue below that productivity growth is the most appropriate measure for capturing the above features.

R&D intensity, defined as the ratio of R&D expenditure to sales, is perhaps less used today than it was in the early empirical IO literature. Intuitively, R&D intensity captures the "innovative effort" of a firm, but it is not a measure of innovative output. The latter ultimately makes R&D intensity an inappropriate measure of innovation. But even as a measure of effort, it suffers from a variety of problems. Notably, there is no requirement for R&D expenditure to be reported, and the peculiarities of firm-level accounting methods make it an unsuitable as an estimate of relative innovative activity (Blundell et al., 1999)

Patent statistics are often used as a measure of innovation because patents represent a novel idea. A citation-weighted patent count, in particular, is the number of forward citations a patent has; its use is based on the idea that more citations are indicative of greater economic value. There are several problems with patents as a measure of innovation, however. One, patents represent only a fraction of innovative output (Griliches, 1990). Two, the incentive to patent encompasses more than just the intent to protect a novel and valuable idea; companies also seek out patents to defend against litigation or to litigate themselves (Hall and Harhoff, 2012). Three, evidence suggests that companies view patents as relatively weak mechanisms for protecting intellectual assets; instead, trade secrecy and lead time to market are viewed as more effective (Cohen et al., 2000). Four, the presumption that a greater number of citations reflect a more valuable idea has been challenged. Empirical evidence indicates that the relationship between citations and economic value is an inverted-U (Abrams et al., 2013). Intuitively, the most valuable patents are the ones firms actively try to protect the most, resulting in fewer citations.

For the reasons above, it has been argued that productivity growth is the most appropriate measure of innovation. Productivity growth, as has been argued by Dale Jorgenson and others at the NBER (see, for example, Jorgenson, 2011), is the key economic indicator of innovation:

"…Productivity growth is the key economic indicator of *innovation*. Economic growth can take place without innovation through *replication* of established technologies. Investment increases the availability of these technologies, while the labor force expands as population grows. With only replication and without innovation, output will increase in proportion to capital and labor inputs, as suggested by Schultz (1956, 1962). By contrast the successful introduction of new products and new or altered processes, organization structures, systems, and business models generates growth of output that exceeds the growth of capital and labor inputs. This results in growth in multifactor productivity or output per unit of input…"

Thus, not only does productivity as a measure avoid the economic complications of patents, it is consistent with long-run economic growth—precisely why Schumpeter was such an advocate for anything that would serve to promote innovation.

With the above in mind, the main concern of the early empirical IO literature was the effect of firm size and market concentration on innovative activity, the latter being some measure of R&D expenditure. Little was found in the way of a statistically robust or economically significant cross-sectional size effect. More recent empirical IO studies have turned their attention to the relationship between *ex-ante* market power and innovation. Typical measures of competition and innovation in this literature are the Lerner index and patent statistics, respectively. The results are mixed despite the use of panel data, similar methods and more sophisticated econometric techniques. Some studies find a negative relationship, some find a positive relationship and some find an inverted-U relationship.

A related but distinct empirical literature presents strong evidence of a positive relationship between competition and innovation. This literature examines changes in productivity at the micro level in response to discrete changes in the competitive environment. In particular, this literature uses natural experiments to identify changes in competitive intensity. For example, one study finds that a group of U.S. iron ore manufacturers nearly doubled their labor productivity in the 1980s after Brazilian manufacturers entered their market.

Economic theory also examines the relationship between competition and innovation. Broadly speaking, two approaches have been taken. One approach compares how much an incumbent firm would be willing to invest in R&D compared to a potential entrant, and the other approach examines the relationship between a parameter that affects market structure and R&D effort. Like the empirical IO literature, the results are mixed. The incumbent-entrant class of models predicts both a positive and negative competition-innovation relationship, while predictions from the parametric class of models range from a positive relationship to an inverted-U.

Overall, both empirical evidence and theory paint an unclear picture of the competition-innovation relationship. The nebulous qualities and complex interaction of these variables is doubtless a major reason why. Our endeavor in this section is to provide a synthesis of the competition-innovation debate and the implications it has for economic growth.

9.1 Early IO Literature: Market Structure, Incumbency, and the Incentive to Innovate

9.1.1 Theory

Arrow (1962) was the first to rigorously examine the relationship between market structure and the incentive to innovate, the latter defined as the difference between post- and pre-innovation profit. Under the crucial assumption that property rights are perfectly exclusive and infinitely lived, Arrow analyzes the relationship by comparing a monopolist's incentive to reduce marginal cost to that of a perfectly competitive firm.

Two types of process innovation are possible in his model: drastic and non-drastic. With drastic innovation, the new technology achieves a profit maximizing price that is less than the marginal cost of the old technology. Consequently, the monopolist remains a monopolist with drastic innovation, and the perfectly competitive firm becomes one.[7] It follows that the incentive to innovate will be higher for a competitive firm since its post-innovation profit is the same as the monopolist's, while its pre-innovation profit of zero is less.

In the case of non-drastic innovation (i.e., the new profit maximizing price is higher than the marginal cost of the old technology), the post-innovation profit for a competitive firm will be less than a monopolist's. This is a consequence of the fact that a competitive firm cannot profitably charge a price that exceeds the prevailing competitive price. Instead, a competitive firm's post-innovation profit will be limited to the unit royalty it charges (i.e., the difference between the old and new marginal costs) times the number of units sold in the market at the old price. Notwithstanding, Arrow shows that a monopolist's incentive to innovate will still be less than a competitive firm's, precisely because the monopolist applies the cost reduction to relatively less units of output in equilibrium.

Gilbert and Newbery (1982) build on Arrow's model by allowing for the possibility of entry into the monopolist's market. Specifically, an incumbent (i.e., monopolist) bids for a patent on a new, substitute technology to preempt entry by a challenger, and a challenger bids for a chance to compete with the incumbent.

In this setup, Gilbert and Newbery find that preemption is a rational strategy—in fact, a Nash equilibrium—for the incumbent if monopoly profits with the new technology exceed the costs of preemption. This will attain if post-entry industry profit is less than pre-emptive monopoly profit. That is, letting e and m denote entrant and monopolist, respectively, if $\pi_m(p_m^1, p_m^2) > \pi_m(p_m^1, p_e^2) + \pi_e(p_m^1, p_e^2)$ attains, where $\pi_i(\cdot)$, $i \in \{m, e\}$, represents the firm's profit function, and p_i^j, $j \in \{1, 2\}$, represents the price of product j, then the incumbent will have the incentive to submit a larger bid than the challenger.[8] Gilbert and Newbery show that this condition holds under fairly weak assumptions, intuitively implying that the incumbent's incentive to innovate will be relatively greater if it has more to lose from entry than the challenger has to gain.

[7] To see the latter, note that the marginal cost of the old technology is equal to the prevailing price in a perfectly competitive market. Thus, with drastic innovation and exclusive property rights, a perfectly competitive firm will drive its competitors out of the market with the new price and subsequently become a monopolist.

[8] Product *j=2* represents the new, substitute technology.

Reinganum (1983) extends the model of Gilbert and Newbery (1982) by introducing uncertainty to the innovation process. Specifically, she assumes that innovation follows a Poisson process with an exponentially distributed date of successful innovation. In turn, the incumbent and challenger choose investment levels that increase their chance of innovating first.

Her model has three possible outcomes. If the incumbent succeeds in reducing its current marginal cost from $\overline{c}$ to $c < \overline{c}$ and secures a patent before the challenger, then it will earn flow profits $\Pi(c)$. If the challenger succeeds before the incumbent, then the incumbent and challenger will earn Cournot flow profits $\pi_I(c) < \Pi(c)$ and $\pi_C(c)$, respectively. If neither succeeds, then the incumbent and challenger maintain their pre-innovation profit flows at R and zero, respectively.[9]

In the spirit of Arrow (1962), Reinganum defines drastic innovation as an innovation that lowers marginal cost to a level $c \leq c_0$, where c_0 is the largest value of c such that $\pi_I(c) = 0$. Coupled with the assumption of constant returns to scale, this condition implies that the incumbent produces zero output. The incumbent is thus knocked out of the market by the challenger when innovation is drastic, allowing the latter to earn profit flows $\pi_C(c) = \Pi(c)$.

The Nash equilibrium of this game implies that the challenger will unambiguously spend more on R&D than the incumbent when innovation is drastic. And to the extent that the equilibrium solution can be analyzed under non-drastic innovation, Reinganum shows that there exists a non-trivial set of non-drastic innovations where the challenger's incentive to innovate is relatively greater.[10]

The assumption of stochastic innovation is critical to the innovation-market structure relationship, as demonstrated by the different results reached by Reinganum (1983) and Gilbert and Newbery (1982). Why the difference occurs may be explained by how investment is modeled. Specifically, whereas in the model of Gilbert and Newbery there is no investment decision (the incumbent and challenger innovate with probability one), Reinganum assumes that further investment can only marginally increase the probability of successful innovation; i.e., the assumptions of Reinganum's model imply diminishing returns to investment. The marginal

---

[9] Reinganum assumes that $\Pi(c)$ and $\pi_C(c)$ are non-increasing and that $\pi_I(c)$ is non-decreasing in c, the intuition being that the successful (unsuccessful) innovator's flow profits are higher (lower) the greater is the reduction in cost.

[10] This is for non-drastic innovations where the reduction in cost is sufficiently close to the drastic level c_0.

increase in expected post-innovation flow profits (via the marginal increase in probability of relatively early innovation) is accordingly offset by the marginal cost of further investment at some point. And this point occurs earlier for the incumbent precisely because its pre-existing profit flows must be replaced.

9.1.2 Empirical analysis

The early empirical literature on firm size, market structure and innovation is voluminous (Gilbert, 2006; Symeonidis, 1996). One paper in particular stands out as an iconic example of this early literature, namely the work of Scherer (1967). Scherer (1967) is actually an extension of Scherer's earlier analysis, Scherer (1965). There are two main differences. First, Scherer (1967) uses a more comprehensive dataset of 56 industries, as opposed to the 48 used in Scherer (1965). Second, whereas Scherer (1965) estimates the impact of firm size on patents granted, Scherer (1967) estimates the effect of market concentration on the total number of engineers and scientists employed as a proportion of total employees. Thus, he moves from an output-based to an input-based measure of innovation.

Two findings stand out in Scherer's cross-sectional study. One is that the explanatory power of market concentration drops substantially when industry-specific dummies are added to the regression. In other words, industry differences in technological opportunity account for most of the variation in R&D investment. Another notable finding is that innovation effort exhibits an inverted-U relationship with concentration. Thus, R&D investment increases with market concentration when market concentration is relatively low, and decreases when it is relatively high. Scherer interprets this as are a rejection of the Schumpeterian hypothesis that market power fosters innovation.

The drop in market concentration's explanatory power after technological opportunity is accounted for is a common theme in the early empirical literature (Symeonidis, 1996). However, it is unclear from these studies if the elimination of market concentration's effect is theoretically driven, or if it is a statistical artifact of the limitations of cross-sectional methods. The latter is addressed by more recent empirical studies, which use panel data methods to identify the effect of market power on innovation. More significant from a policy perspective is the inverted-U relationship brought forward by Scherer, which implies that too much competitive rivalry will retard innovation. The possibility of a non-linear competition-innovation relationship has been and continues to be investigated in the IO literature. We discuss this more recent literature shortly.

9.2 New IO Literature: Competition as a Parameter and the Incentive to Innovate

9.2.1 Theory

Early theoretical models may be viewed as "discrete" in the sense that they compare the incentive to innovate across two types of firms: one with market power and one without. "New" theory, however, examines the competition-innovation relationship with competition measured on a continuum. Kamien and Schwartz (1976) were the first to take this approach.

In their model, a firm chooses a date to introduce innovation based on the degree of rivalry it faces, where greater rivalry is modeled as a parameter that accelerates the expected date of rival innovation. Specifically, the decision problem facing the firm is to choose an innovation arrival time that maximizes its present value of expected innovation cash flows, conditional on the probability of a rival innovating first.[11] The firm formally chooses its development date T^* to solve

$$\max_T \int_T^{\infty} e^{-(r-g)t}[P_0(1-F(t)) + P_1(F(t)-F(T)) + P_2F(T)]dt - C(T),$$

where $r > g$ is the discount rate; g is market growth; $F(\tau)$ is the probability of rival introduction by time τ; $P_0(1-F(t))$ is the expected payoff to the firm conditional on no rival entry by time t; $P_1(F(t)-F(T))$ is the expected payoff to the firm conditional on it innovating before a rival and the rival appearing between time T and t; P_2 is the expected payoff to the firm conditional on a rival innovating before time T; and $C(T)$ is the minimum present value of the cost of completing development by time T. It is assumed that $C'(T) < 0$ and $C''(T) > 0$, and $P_0 \geq P_1$ and $P_0 \geq P_2$.

In addition to other parameters, the optimal date T^* is a function of the hazard rate of successful innovation by a rival, $h \equiv F'(t)/(1-F(t))$. Kamien and Schwartz assume for simplicity that the probability of innovation is memoryless, implying a constant hazard rate with cumulative density function $F(\tau) = 1 - e^{h\tau}$, $\tau \in [0, T]$. The significance of h is its tie to the degree

[11] Kamien and Schwartz (1976) examine two versions of this model, one with patent protection and one without. We omit the model with patent protection as the results are not affected by the differences in appropriability.

of rivalry. Namely, because the inverse of h is the expected date of innovation by a rival, an increase in h heightens the probability of a rival innovating first. Kamien and Schwartz accordingly interpret an increase in h as an intensification of rivalry.

As Kamien and Schwartz demonstrate, the relationship between h and T^* depends on the size of the innovation (defined as the amount of profit rewarded for innovating). When the size of innovation is sufficiently large, the relationship takes a U shape. That is, for relatively small values of h, the marginal cost saving of postponement is less than the marginal loss of delay, pushing the firm to accelerate its development date. But for sufficiently high values of h, the effort required to pre-empt a rival's development date becomes excessively costly, driving the firm to postpone. In the case of small innovation size, h and T^* are strictly positively related. Intuitively, an already small reward for innovation makes any additional effort to pre-empt rival innovation less attractive.

Loury (1979) extends the model of Kamien and Schwartz (1976) to a game-theoretic setting. In particular, he examines how R&D incentives are affected by rivalry when a finite number of symmetric firms, n, choose their investment strategies simultaneously. As in Kamien and Schwartz (1976), Loury assumes that the date of successful innovation is memoryless. Thus, the probability of successful innovation by time t is given by $P(\tau(x) \leq t) = 1 - e^{-h(x)t}$, where $\tau(\cdot)$ is the random date of successful innovation, x is R&D expenditure and $h(\cdot)$ is the hazard rate of innovation. It is assumed that $h'(x) > 0$, $h''(x) \geq 0$ for $x \in [0, x]$ and $h''(x) < 0$ for $x \in (x, \infty)$.

Each firm chooses its R&D expenditure to maximize its present value of discounted cash flows, taking as given the collective R&D expenditure of its rivals. The symmetry of the problem implicitly defines the optimal R&D expenditure function $x^* = \hat{x}(a, r, V)$, where $a = (n-1)h(x^*)$ and V and r denote the flow of revenues from successful innovation and the interest rate, respectively.

In the special case of no strategic effects, Loury shows that a firm's R&D expenditure will either increase then decrease with rival R&D expenditure (i.e., the parameter a), or just decrease. This mirrors the results of Kamien and Schwartz (1976). When strategic effects are allowed (and $n \geq 2$), however, individual R&D expenditure and the expected date of innovation strictly decrease with n. Hence, while greater rivalry (as measured by an increase in n) lowers individual R&D expenditures, it also accelerates the introduction of innovation. Intuitively, an increase in the number of firms in a symmetric industry will lower expected profits, thus deterring firms to invest in R&D; concurrently, more firms will simultaneously invest in R&D, which increases the chance of earlier innovation.

Aghion et al. (2001) (henceforth, AHHV) develop a dynamic macroeconomic model of "step-by-step" innovation whereby a leading (laggard) firm in its sector can widen (narrow) its technological lead (lag) if it successfully innovates. The continuum of sectors that makeup the economy are each comprised of two cost-asymmetric firms that ultimately compete in price with differentiated goods.

In the first stage of a two-stage game, each firm takes their second-stage profit function as given and makes an R&D investment decision to maximize expected future discounted profits, conditional on the technological gap in their industry and the level of competition. In particular, the firm makes a cost-reducing investment that increases its probability of favorably changing the technological gap in its industry, where the technological gap is a function of the relative production cost between the two firms. In the second stage, firms compete on price alone, taking as given their relative cost. The equilibrium profit in this stage is an implicit function of relative marginal cost[12] and the degree of product substitutability. The latter is AHHV's measure of competition; however, it is more precisely a reflection of consumer tastes, which affects the interpretation of an industry's market structure.[13]

AHHV analyze the effect of competition on innovation under the following assumptions: (1) there is a steady state composition of *n*-gap (unleveled) and 0-gap (leveled) industries; (2) the laggard can immediately catch-up to the leader if the laggard is the sole innovator; (3) there is at-most a one-step increase in the gap if a leader or neck-and-neck firm innovates and; (4) changes in product substitutability affect the whole economy.

When innovation is large, they find that an increase in competition can either stimulate or retard innovation. This is for two reasons. First, large innovation implies that a one-step lead will raise the would-be leader's profit to the maximal level, so a leader will not innovate further.

[12] Defined as the ratio of firm *i*'s marginal cost to firm *j*'s marginal cost.

[13] AHHV concede that product substitutability is a taste parameter. By construction, product substitutability will affect the structure of a firm's demand. When product substitutability is at its lowest, a firm has no competitors, so the firm will behave like a monopolist. On the other hand, a firm will behave like a perfectly competitive firm when product substitutability is at its highest. Thus, a higher level of product substitutability may be interpreted as a "less monopolistic" market structure.

Second, large innovation effectively fixes the post-innovation rent for a neck-and-neck firm and the pre-innovation rent for a laggard. The effect of competition on innovation thus only operates through the pre- and post-innovation rents of neck-and-neck and laggard firms, respectively. In fact, these rents are the same: the pre-innovation rent for a neck-and-neck firm is the post-innovation rent for a laggard. Thus, since greater competition lowers neck-and-neck rents, competition will have a non-linear effect on economy-wide innovation.

The shape of the relationship found by AHHV under large innovation is an inverted-U. This arises because the steady state distribution of leveled and unleveled industries is itself a function of the level of competition. More specifically, when competition is already relatively high, neck-and-neck firms will have a greater incentive to innovate than laggards as they seek to "escape competition." This will push the economy into a state where there are more unleveled industries than leveled ones. And because this transition does not affect the level of competition (i.e., the level of neck-and-neck profits), the incentive to innovate across the economy will diminish—laggards have little incentive to innovate when competition is already high, and leaders do not innovate at all. AHHV call this the "Schumpeterian" effect of competition. Contrast this to the case of relatively low competition. When competition is relatively low, laggard firms will be more inclined than neck-and-neck firms to innovate. This will push the economy into a state with more leveled than unleveled industries where the "escape competition" effect dominates. Thus, innovation initially increases with competition but then declines.

In the case of small innovation, innovation is found to monotonically increase with competition. Intuitively, the increment in profit from innovating is approximately the same for leaders, laggards and neck-and-neck firms; and since the increment in profit for a neck-and-neck firm increases with competition (due to the "escape competition" effect), it follows that economy-wide innovation will also increase.

Finally, AHHV examine the general case numerically and confirm their analytical results. That is, they find innovation to increase with competition for intermediate values of innovation size, but eventually an inverted-U shape arises when innovation is large and the probability of imitation is low.

Aghion et al. (2005) (henceforth, ABBGH) reexamine the "step-by-step" innovation model developed in AHHV under two modifications. First, a firm can only advance its technological position by one step through successful innovation. If, for example, the current state of a sector is m—a non-negative integer that indexes the efficiency gap between two firms—and the leader (laggard) successfully innovates while the laggard (leader) does not, then the state of the sector will change from m to $m + 1$ ($m - 1$). Thus, unlike AHHV, this model assumes that the laggard cannot immediately catch up to the leader.

Second, instead of explicitly using product substitutability as a proxy for competition, greater competition is measured as the degree to which neck-and-neck firms cannot collude. That is, letting $\epsilon \in [0, 1/2]$ denote a neck-and-neck firm's profit as a fraction of a leader's profit, ABBGH treat a smaller value of ϵ as greater competition. To operationalize ϵ in their model, ABBGH use $\Delta \equiv 1 - \epsilon = (\pi_1 - \pi_0)/\pi_1$, where π_0 and π_1 are the profit levels of a neck-and-neck and leader firm, respectively.

With the above modifications, ABBGH restrict attention to the case where $m \in \{0, 1\}$, so that a leader (laggard) can be at most one step ahead (behind). This mirrors the "large innovation" scenario analyzed in AHHV, and implies that a leader will not innovate. The inverted-U competition-innovation relationship found in AHHV (under large innovation) remains intact in ABBGH.

Hashmi (2013) argues that the theoretical model of ABBGH is not well-suited for industry level analysis, particularly because ABBGH model the interaction of competition and innovation at the economy level. Hashmi notes, however, that an industry-level analysis is possible with ABBGH's duopoly model of competition, so he accordingly adopts it.

Namely, Hashmi considers a setting where two cost-asymmetric firms price compete with differentiated goods. The demand for firm i, $i = 1, 2$, is given by:

$$q_i = \frac{p_i^{\frac{1}{\alpha-1}}}{p_i^{\frac{\alpha}{\alpha-1}} + p_j^{\frac{\alpha}{\alpha-1}}},$$

where p is price, and $\alpha \in [0, 1]$ is the degree of product substitutability. Hashmi follows ABBGH and uses α to measure competition. Letting $c_i = w\gamma^{-k_i}$ denote firm i's constant marginal cost of production, where w is the wage rate, γ is the size of innovation and k_i is the technology level of firm i, the equilibrium profit function for firm i is

$$\pi_i(n) = \frac{(1-\alpha)R_i(n)}{1-\alpha R_i(n)},$$

where $n \equiv k_i - k_j$ is the technology gap between firm i and firm j, and $R_i(\cdot)$ is firm i's market share. Finally, Hashmi defines the probability of successful innovation as

$$P = [1 - e^{-ax}] + \max\{0, 1 - e^{\eta(n-\hat{n})}\},$$

where x is R&D investment. The first term is common to both firms; it represents the baseline probability of success. The second term is specific to an unleveled industry, and it allows the laggard to more easily catch up to the leader as the technological gap grows. This term is introduced by Hashmi to ensure that zero investment is not chosen by both the laggard and leader when the gap is large.

Given the above, both firms choose their R&D investment level to maximize their expected future discounted profits. The relationship between optimal R&D investment and competition depends on the technology gap. When the technology gap is small, the relationship is approximately monotonically increasing; when it is intermediate, the relationship is an inverted-U; and when it is large, the relationship is monotonically decreasing. Thus, ABBGH's result is not robust to variations in the technology gap.

Hashmi's results are interesting from a policy perspective, particularly if one views the technology gap as a better proxy, or at least a characteristic, of market structure. Intuitively, a technological lead confers a competitive advantage (in this case, the ability to charge a more competitive price due to a cost-reducing technology), and as the lead gets larger, the industry will move closer to monopoly.[14] Perhaps even more interesting is the result that innovation decreases with the technological gap. Ostensibly, a large gap lulls the leader and discourages the

[14] In fact, when an industry consists only of two firms, a larger technological gap will necessarily lead to greater concentration. This relationship does not necessarily hold for an industry with more than two firms, however. When there are three firms with different efficiency levels, for example, the laggard firm could narrow the technological gap with the second-place and leader firm, but the reallocation of sales among the three firms could lead to lower or higher concentration.

laggard from innovating, whereas a small gap keeps rivals from becoming complacent with their competitive positions.

9.2.2 Empirical analyses

Early empirical IO studies were directly interested in the hypothesis that large scale stimulates innovation. More recent empirical IO research is expressly concerned with the effect of competition on innovation. In addition to this different, albeit somewhat related policy question, the newer literature differentiates itself via more sophisticated econometric techniques, measures of competition and innovation, and breadth of data.

The literature is broadly split into two categories; models that test competition's effect on innovation under the assumption that the relationship is linear, and models that test the effect assuming the relationship is non-linear. In particular, Nickell (1996) and Blundell et al. (1999) estimate linear specifications, while ABBGH, Correa (2012), Correa and Ornaghi (2014), and Hashmi (2013) estimate non-linear specifications.

Both Nickell (1996) and Blundell et al. (1999) estimate a linear, dynamic panel-data model and find a positive competition-innovation relationship based on a panel of publicly traded firms in the UK. Nickell uses a firm's average operating margin as a proxy for competition and total factor productivity to measure of innovation, whereas Blundell et al. (1999) measure competition in terms of individual market share, market concentration and import penetration, and innovation in terms of survey-based innovation counts.

The non-linear class of empirical IO models, all based on ABBGH, generate a range of qualitative results. ABBGH use Poisson regression and 1973-1994 panel of publicly traded firms in the UK to estimate the competition-innovation relationship at the industry-year level. They use citation-weighted patent counts to measure innovation and operating margins to measure competition. They find that competition's effect on innovation is positive for relatively low and intermediate levels of competition, but negative when competition is sufficiently high. Using the same data and econometric techniques as ABBGH, Correa (2012) instead finds a monotonically increasing relationship after accounting for structural changes in the data, namely a policy that facilitated patent grants. Correa and Ornaghi (2014) also find a monotonically increasing relationship, but with panel data on US manufacturing firms spanning the period 1974-2001.

Their finding is robust to different measures of innovation, including citation-weighted patent counts and productivity growth. Finally, Hashmi (2013) finds a weak, monotonically negative relationship using a panel of US publicly traded firms over the period 1976-2001. His econometric model is largely the same as ABBGH's, except he uses negative binomial regression to estimate the competition-innovation relationship. The range of conclusions for the non-linear class of models is striking given the strong similarity in data and methods. This bears further discussion, which we pursue below.

As noted above, the non-linear empirical models that test the competition-innovation relationship are closely related to the ABBGH model. Accordingly, this warrants some discussion of the ABBGH model and the models that followed.

Specifically, ABBGH estimate following conditional mean function:

$$\mathrm{E}[p|c,x] = e^{g(c)+x\beta},$$

where p is the number of citation-weighted patents, c is competition, x is a set of industry and time dummy variables and $g(\cdot)$ is some function to be estimated. It is assumed that p follows a Poisson process. ABBGH estimate the model with data on publicly traded manufacturing firms in the UK over the period 1973 to 1994.

Correa (2012) revisits the ABBGH model using the same sample of data and empirical formulation, but allows for structural breaks in the data. Correa reasons that this is appropriate because the establishment of the United States Court of Appeals for the Federal Circuit (henceforth CAFC) in 1982 made it effectively easier to have a patent granted. Correa takes two approaches to test the structural break hypothesis. The first approach is a Chow test. The base model estimated is the same as in ABBGH:

$$p_{jt} = \exp\{\beta_0 + \beta_1 c_{jt} + \beta_2 c_{jt}^2 + \varphi \hat{v}_{jt} + \delta_1 D_\tau c_{jt} + \delta_2 D_\tau c_{jt}^2 + \textstyle\sum_j \alpha_j D_j + \sum_t \gamma_t D_t + u_{jt}\},$$

where $D_\tau = 1$ for all $t \geq \pi$ (π denotes a pre-defined structural break), 0 otherwise; c_{jt} is the level of competition for industry j at time t, measured as one minus the industry average price-cost margin; $\hat{v}_{jt}$ is the residual for industry j at time t from regressing the competition index on policy and foreign-industry instruments (i.e., endogeneity is accounted for with a control function approach); and the last two terms are industry- and time-fixed effects. The null hypothesis of time stability at $t = 1983$ is rejected by the Chow test at the 5% level of significance.

Correa also carries out a Wald-type test for structural breaks, finding only one structural break at year 1981. Correa gives several reasons why this year, instead of 1982, was detected. One of these reasons is that the political discussion to establish the CAFC began in 1979. Thus, in anticipation of the CAFC being established, a structural change in patent incentives may have manifested before the CAFC's official introduction.

The indication of a structural break in the data by both tests prompted Correa to test the joint statistical significance of the competition coefficients. Correa concluded from this exercise that before the establishment of the CAFC, the competition-innovation relationship is statistically significant; but after the CAFC's introduction, the relationship is not statistically significant. Correa then estimated the relationship under the two regimes, i.e., with the two identified structural breaks. In both cases, he found that the relationship between competition and innovation is monotonically increasing before the structural break.

Hashmi (2013) and Correa and Ornaghi (2014) also revisit the empirical ABBGH model. Both use a negative binomial instead of a Poisson specification for the conditional mean function (to account for over-dispersion in the data), data on publicly traded manufacturing firms in the US and the Lerner index as a proxy for competition. Nevertheless, Correa and Ornaghi (2014) find a monotonically positive relationship using total factor productivity growth, labor productivity and citation-weighted patent counts as a proxy for innovation, while Hashmi (2013)|, using only citation-weighted patent counts as a proxy for innovation, finds a monotonically "mild," but negative relationship.

9.2.2 Using natural experiments to identify the effect of competition

The empirical IO literature on competition and innovation has exclusively focused on concentration and profitability to infer the level of competitiveness of an industry, and largely patents and R&D to capture the level of innovation. Another literature investigates the link between structural changes to the competitive environment and productivity. Following Holmes and Schmitz (2010), this section presents a brief review of that literature.

We would first like to draw attention to some comments made by Holmes and Schmitz (2010) regarding the measurement of competition and how competition affects productivity. Holmes and Schmitz (2010) claim that concentration and profitability are inadequate at identifying structural changes in a competitive environment and, in fact, have the potential to mislead. To illustrate, they consider an industry that has a strong trade barrier and is made up of small, unproductive firms. The government then lifts the trade barrier, subsequently drawing the attention of large, highly productive to enter the market. From the perspective of the researcher who observes only market shares and profitability, they might conclude that the industry became less competitive due to a significant increase in concentration and profitability. This, however, is at odds with the conventional thinking that less entry barriers stimulate competition. They also note the selection effect of competition, whereby relatively unproductive firms are “selected out” of an industry because they cannot compete effectively. Taken altogether, market concentration may *increase* in a more competitive state. Moreover, to the extent that productivity is positively correlated with profitability, average profitability will also *increase* in a more competitive state.

Notwithstanding the above, Holmes and Schmitz (2010) note that there is no model that can comprehensively explain why or how competitive pressure induces firms to be more productive. The body of evidence strongly suggests that it does, however. Matsa (2011), for example, found that incumbent supermarket retailers significantly upgraded their inventory systems (to maximize product availability) after Wal-Mart entered their local markets. Importantly, the increase in productivity cannot be attributed to an increase in market or average firm sizeas demand did not all of a sudden increase for the existing firms, nor did the existing firms substantially increase in scale. The observed gain in productivity was, therefore, largely a response to increased competitive pressure.

Competition can also boost productivity by lowering the opportunity cost of investment. For example, Schmitz (2005) found that plant managers were reluctant to adopt new managerial practices because they feared losing profits to a job strike. From this, he argues that the competitive process—which in the absence of innovation tends to shrink margins over time—will reduce forgone profits and thereby spur investment into new, more efficient forms of management.

Major shifts in the competitive landscape are perhaps the best way to identify the effect of competition on innovation. Holmes and Schmitz (2001) examine the effect of railroad transportation on water shipping in the 19th and 20th centuries. Before transportation by railroad was economically feasible in the US (1850s), freight transportation by water was effectively the only way to ship cargo across the nation. This meant that ports not only had tremendous market power, but also had the incentive to keep it. Their market power was heavily weakened, however, when railroads became a viable alternative for transportation.

Railroads undermined the market power of ports in two ways. On one hand, railroads gave consumers easier accessibility to other ports. So, if a consumer was not happy with the price or service of a port, it could use a train to ship its cargo to another port. On the other hand, railroads could in some cases entirely replace the function provided by water services. Ultimately, the threat that railroads presented to ports manifested as an effort by the latter to increase productivity.

In the same vein, Galdon-Sanchez and Schmitz (2002) and Schmitz (2005) examine the effect of Brazil's entry into the lower Great Lakes iron ore market during the 1980s. They note that before Brazil's entry markets were characterized by few and distant producing locations and high transportation costs, which fostered significant market power. Market power was in fact evident since unions and local government exercised their power to extract as much of the surplus from iron ore producers as possible.

The market power of iron ore producers around the Great Lakes eventually eroded, however. Due to a substantial decrease in transportation costs, Brazil entered the iron market around the Great Lakes in the 1980s. This put tremendous price pressure on the domestic iron ore producers and, in turn, pressure to improve labor productivity. Labor productivity actually doubled in the mid-1980s, and Galdon-Sanchez and Schmitz (2002) demonstrate that the source of productivity growth was not due to the closing of inefficient mines or increases in scale, but rather surviving mines that made investments to lower costs.

Finally, Syverson (2004) investigates the effect of spatially dense competition on the distribution of productivity in the US ready-mix concrete industry. He finds that more densely clustered markets exhibit higher average productivity and lower productivity dispersion. The reason for this is that more densely clustered markets lower switching costs for consumers, and thus an inefficient firm is more likely to exit an industry that is highly dense. In turn, average productivity and productivity dispersion will increase and decrease, respectively, in more competitive markets.

9.2.2 Concluding Remarks

In this Chapter, we have demonstrated what market power is, how it is measured and how it can impact productivity and innovation. Such a chapter of course leaves much room for other studies and perspectives and we do not expect that this Handbook chapter will satisfy all of those in this rather dense literature. We do trust, however, that our treatments and perspectives on important contributions to this literature are balanced and that they provide a relatively complete perspective on such an important issue in regard to a broad array of assumptions and methodological approaches.

10. References

Abrams, D., Akcigit, U., and Popadak, J. (2013), Patent value and citations: Creative destruction or strategic disruption?, NBER working paper, 19647

Aghion, P., Harris, C., Howitt, P., and Vickers, J. (2001), Competition, imitation and growth with step-by-step innovation, *Review of Economic Studies*, 68, 467-492.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005), Competition and innovation: an inverted-u relationship, *Quarterly Journal of Economics*, 120, 701-728.

Appelbaum, E. (1982), The estimation of the degree of oligopoly power, *Journal of Econometrics*, 19, 287-299.

Armstrong, M. (2006), Recent developments in the economics of price discrimination, Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress of the Econometric Society, ed. by R. Blundell, W. Newey, and T. Persson. Cambridge University Press, Cambridge, UK.

Arrow, K.J. (1962), Economic welfare and the allocation of resources for invention. In The Rate and Direction of Inventive Activity (R. Nelson, ed.), 609-625, Princeton University Press, Princeton, NJ.

Bain, J.S. (1956), Barriers to new competition, Harvard University Press. Cambridge, MA.

Berry, S., Levinsohn, J., and Pakes, A. (1995), Automobile prices in market equilibrium, *Econometrica*, 63, 841-90.

Bingaman, Anne (1994), Innovation and antitrust. https://www.justice.gov/atr/speech/innovation-and-antitrust.

Blundell, Richard, Griffith, R., and Reenen, J.V.(1999), Market share, market value and innovation in a panel of British manufacturing firms, *Review of Economic Studies*, 66, 529-554.

Boone, J. (2008a), A new way to measure competition, *Economic Journal*, 118, 1245-1261.

Boone, J. (2008b), Competition: theoretical parameterizations and empirical measures, *Journal of Institutional and Theoretical Economics*, 164, 587-611.

Bresnahan, T.F. (1981), Departures from marginal-cost pricing in the American automobile industry, *Journal of Econometrics* 17, 201-227.

Bresnahan, T.F. (1982), The oligopoly solution concept is identified, *Economics Letters*, 10, 87-92.

Bresnahan, T.F. (1987), Competition and collusion in the American automobile oligopoly: The 1955 price war, *Journal of Industrial Economics,* 35, 457-482.

Bresnahan, T.F. (1989), Empirical studies of industries with market power. In: Schmalensee, Richard, Willig, Robert D. (Eds.), The Handbook of Industrial Organization, vol. 2. North-Holland, Amsterdam, 1011-1057.

Bresnahan, T.F. and Reiss, P.C. (1990), Entry in monopoly markets, *Review of Economic Studies*, 57, 57-81.

Brown, B. (1983), The identification problem in systems nonlinear in the variables, *Econometrica*, 51, 175-96.

Brown, D.J. and Matzkin, R.L. (1998), Estimation of nonparametric functions in simultaneous equations models with application to consumer demand, Working paper, Yale University.

Bulow, J. and Klemperer, P. (1999), Prices and the winner's curse, *RAND Journal of Economics*, 33, 1-21.

Cohen, W.M., Nelson, R.R., and Walsh, J.P., Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not), NBER working paper, 7552

Correa, J.A. (2012), Innovation and competition: an unstable relationship, *Journal of Applied Econometrics*, 27, 160-166.

Correa, J.A. and Carmine, O. (2014), Competition and innovation: evidence from U.S. patent and productivity data, *Journal of Industrial Economics*, LXII, 258-285.

Corts, K.S. (1999), Conduct parameters and the measurement of market power, *Journal of Econometrics*, 88, 227-250.

Deaton, A. and Muellbauer, J.N. (1980), Economics and Consumer Behavior. New York: Cambridge University Press.

Demsetz, H. (1973), Industry structure, market rivalry, and public policy, *Journal of Law and Economics*, 16, 1-9.

Driskill, R. and McCafferty, S. (1989), Dynamic duopoly with adjustment costs: A differential game approach, *Journal of Economic Theory*, 49, 324-38.

Ellickson, P. (2007), Does Sutton apply to Supermarkets?, *RAND Journal of Economics,* 38, 43-59.

Ellickson, P. (2015), Market structure and performance, in James D. Wright (Ed.), International Encyclopedia of the Social and Behavioral Sciences. 2nd Edition, Vol. 14: 9211-9216 (Oxford, Elsevier).

Fisher, F.M. and McGowan, J.J. (1983), On the misuse of accounting rate of return to infer monopoly profits, *American Economic Reviews*, 73, 1141-1200.

Fudenberg, D. and Tirole, J., Game Theory, MIT Press, 1991.

Gabszewicz, J. and Thisse, J.F. (1979), Price competition, quality and income disparities, *Journal of Economic Theory*, 20, 340-359.

Gabszewicz, J. and Thisse, J.F. (1980), Entry, (and exit) in a differentiated industry, *Journal of Economic Theory*, 22, 327-338.

Galdon-Sanchez, J.E. and James A.S. (2002), Competitive pressure and labor productivity: world iron-ore markets in the 1980s, *American Economic Review*, 94, 1222-1235.

Garcia, D.D. (2016) Examination of the relationship between competition and innovation: Toward a robust approach, unpublished PhD dissertation, Rice University.

Gasmi, F., Laffont, J.J., and Vuong, Q.H. (1992), Econometric analysis of collusive behavior in a soft-drink market, *Journal of Economics and Management Strategy,* 1, 277-311.

Gilbert, Richard (2006), Looking for Mr. Schumpeter: where are we in the competition-innovation debate? In *Innovation Policy and the Economy* (Adam B. Jaffe, Josh Lerner, and Scott Stern, eds.), 159-215, The MIT Press.

Gilbert, R.J. and David M.N. (1982), Preemptive patenting and the persistence of monopoly, *American Economic Review*, 72, 514-526.

Gollop, F.M., and Roberts, M.J. (1979), Firm interdependence in oligopolistic markets, *Journal of Econometrics,* 16, 617-45.

Griliches, Zvi (1990), Patent statistics as indicators: A survey, *Journal of Economic Literature*, XXVIII, 1661-1707.

Graddy, K. (1995), Testing for imperfect competition at the Fulton fish market, *RAND Journal of Economics*, 25, 37-57.

Hall, Bronwyn and Harhoff D., Recent research on the economics of patents, *Annual Review of Economics*, 4, 541-565.

Hashmi, A.R. (2013), Competition and innovation: the inverted-u relationship revisited, *Review of Economics and Statistics*, 95, 1653-1668.

Hausman, J.A., and Gregory K.L. (2004), The competitive effects of a new product introduction: A case study, *Journal of Industrial Economics*, 50, 237-63.

Hazledine, T. (2006), Price discrimination in Cournot-Nash oligopoly, *Economics Letters*, 93, 413-420.

Hazledine, T. (2010), Oligopoly price discrimination with many prices, *Economics Letters*, 109, 150-153.

Hashmi, A.R. (2013), Competition and innovation: the inverted-u relationship revisited, *Review of Economics and Statistics*, 95, 1653-1668.

HMG (2010), Horizontal merger guidelines: U.S. Department of Justice and Federal Trade Commission. https://www.justice.gov/atr/horizontal-merger-guidelines-08192010.

Holmes, Thomas and James Schmitz (2010), Competition and productivity: a review of evidence. Federal Reserve Bank of Minneapolis: Research Department Staff Report 439.

Hotelling, H. (1929), Stability in competition, *Economic Journal*, 39, 41-57.

Iwata, G. (1974), Measurement of conjectural variations in oligopoly, *Econometrica*, 42, 947-66.

Jorgenson, D.W. (2011), Innovation and productivity growth, *American Journal of Agricultural Economics*, 93, 276-296.

Kamien, M.I. and Nancy, L.S. (1976), On the degree of rivalry for maximum innovative activity, *Quarterly Journal of Economics*, 90, 245-260.

Karakaplan, M.U. and Kutlu, L. (2017), Estimating market power using a composed error model, mimeo.

Klemperer, P. and Meyer, M. (1989), Supply function equilibria in oligopoly under uncertainty, *Econometrica*, 57, 1243-1277.

Kreps, D. and Scheinkman, J. (1983), Quantity precommittment with Bertrand competition, *Bell Journal of Economics*, 14, 326-337.

Kumbhakar, S.C., Baardsen, S., and Lien, G. (2012), A new method for estimating market power with an application to norwegian sawmilling, *Review of Industrial Organization*, 40, 109-129.

Kutlu, L. (2009), Price discrimination in Stackelberg competition, *Journal of Industrial Economics*, 57, 364.

Kutlu, L. (2012), Price discrimination in Cournot competition, *Economics Letters*, 117, 540-543.

Kutlu, L. and Sickles, R.C. (2012), Estimation of market power in the presence of firm level inefficiencies, *Journal of Econometrics*, 168, 141-155.

Kutlu, L. and Wang, R. (2018), Estimation of cost efficiency without cost data, Forthcoming in

Journal of Productivity Analysis.

LaFrance, J.T. (1990), Incomplete demand systems and semilogarithmic demand models, *Australian Journal of Agricultural Economics,* 34, 118-31.

LaFrance, J.T. (2004), Integrability of the linear approximate almost ideal demand system, *Economics Letters,* 84, 297-303.

Lau, L.J. (1982), On identifying the degree of competitiveness from industry price and output Data, *Economics Letters*, 10, 93-99.

Loury, G.C. (1979), Market structure and innovation, *Quarterly Journal of Economics*, 93, 395-410.

Marin, P. and Siotis, G. (2007), Innovation and market structure: An empirical evaluation of the 'bounds approach' in the chemical industry*, Journal of Industrial Economics*, 55, 93-111.

Matsa, D.A. (2011), Competition and product quality in the supermarket industry, *Quarterly Journal of Economics*, 126, 1539-1591.

Nickell, S.J. (1996), Competition and corporate performance, *Journal of Political Economy*, 104, 724-746.

Nevo, A. (2000a), A practitioner's guide to estimation of random-coefficients logit models of demand, *Journal of Economics and Management Strategy*, 9, 513-48.

Nevo, A. (2000b), Mergers with differentiated products: The case of the ready-to-eat cereal industry, *RAND Journal of Economics* 31, 395-421.

Nevo, A. (2001), Measuring market power in the ready-to-eat cereal industry, *Econometrica*, 69, 307-42.

Nocke, V., (2007), Collusion and dynamic (under)investment in quality, *RAND Journal of Economics*, 38, 227-249.

Orea, L. and Steinbuks, J. (2018), Estimating market power in homogenous product markets using a composed error model: Application to the California electricity market, *Economic Inquiry*, 56, 1296-1321..

Perloff, J.M., Karp, L.S., and Golan, A. (2007), Estimating market power and strategies. Cambridge University Press, Cambridge.

Perloff, J.M. and Shen, E.Z. (2012), Collinearity in linear structural models of market power, *Review of Industrial Organization*, 2, 131-138.

Pindyck, R. (1985), The measurement of monopoly power in dynamic markets, *Journal of Law and Economics*, 28, 193-222.

Porter, R. (1983), A study of cartel stability: The joint executive committee 1980–1986, *Bell Journal of Economics*, 14, 301-14.

Puller, S.L. (2007), Pricing and firm conduct in California's deregulated electricity market, *Review of Economics and Statistics*, 89, 75-87.

Puller, S.L. (2009), Estimation of competitive conduct when firms are efficiently colluding: addressing the Corts critique, *Applied Economics Letters*, 16, 1497-1500.

Reinganum, J.F. (1983), Uncertain innovation and the persistence of monopoly, *American Economic Review*, 73, 741-748.

Reynolds, S. (1987), Capacity investment, preemption and commitment, *International Economic Review*, 28, 69-88.

Roehrig, C. S. (1988), Conditions for identification in nonparametric and parametic models, *Econometrica*, 56, 433-47.

Rosenthal, R. (1980), A model in which an increase in the number of sellers leads to a higher price, *Econometrica*, 48, 1575–1579.

Salop, S. (1979), Monopolistic competition with outside goods, *Bell Journal of Economics*, 10, 141-156.

Scherer, F.M. (1965), Firm size, market structure, opportunity, and the output of patented inventions, *American Economic Review*, 55, 1097-1125.

Scherer, F.M. (1967), Market structure and the employment of scientists and engineers, *American Economic Review*, 57, 524-531.

Schmalensee, R. (1989), Inter-industry differences of structure and performance, In: Schmalensee, Richard, Willig, Robert D. (Eds.), The Handbook of Industrial Organization, vol. 2. North-Holland, Amsterdam, 951-1009.

Schmitz, J.A. (2005), What determines productivity? lessons from the dramatic recovery of the U.S. and Canadian iron ore industries following their early 1980s crisis, *Journal of Political Economy*, 113, 582-625.

Schumpeter, J.A. (1942), Capitalism, socialism and democracy. New York: Harper & Row.

Shaked, A. and Sutton, J. (1982), Relaxing price competition through product differentiation, *Review of Economic Studies*, 49, 3-13.

Shapiro, C. (2012), Competition and innovation: did Arrow hit the bull's eye? In The Rate and Direction of Inventive Activity Revisited (Josh Lerner and Scott Stern, eds.), 361-404, University of Chicago Press.

Shaked, A. and Sutton, J. (1983), Natural oligopolies, *Econometrica*, 51, 1469-1484.

Spiller, P.T. and Favaro, E. (1984), The effects of entry regulation or oligopolistic interaction: The Uruguayan banking sector, *RAND Journal of Economics*, 15, 244-54.

Stiglitz, J. (1989), Imperfect information in the product market. In: Schmalensee, Richard, Willig, Robert D. (Eds.), The Handbook of Industrial Organization, vol. 1. North-Holland, Amsterdam, 769-847.

Stole, L.A. (2007), Price discrimination and competition, In: M. Armstrong and R. Porter (Eds.), The Handbook of Industrial Organization, vol. 3, North-Holland, Amsterdam, 2221-2299.

Stone, J.R.N. (1953), The measurement of consumers' expenditure and behavior in the United Kingdom, 1920-38. Cambridge, UK: Cambridge University Press.

Sutton, J. (1991), Sunk Costs and Market Structure, Cambridge, MA: MIT Press.

Sutton, J. (1998), Technology and Market Structure, Cambridge, MA: MIT Press.

Sutton, J. (2007), Market structure: Theory and evidence. In: Armstrong, M. Porter, R. (Eds.) The Handbook of Industrial Organization, vol. 3. North-Holland, Amsterdam, 2301-2368.

Symeonidis, G. (1996), Innovation, firm size and market structure: Schumpeterian hypotheses and some new themes, OECD economics department working paper, 161

Symeonidis, G. (2000), Price competition and market structure: The impact of restrictive practices legislation on concentration in the U.K., *Journal of Industrial Economics*, 48, 1-26.

Symeonidis, G. (2001), The effects of competition: Cartel policy and the evolution of strategy and structure in British industry, Cambridge MA: MIT Press.

Syverson, C. (2004), Market structure and productivity: a concrete example, *Journal of Political Economy*, 112, 1181-1222.

Utterback, J. and William, A. (1975), A dynamic model of process and product innovation, *Omega*, 3, 639-656.

Vuong, Q.H. (1989), Likelihood ratio tests for model selection and non-nested hypotheses, *Econometrica*, 57, 307-33.

Weiss, L. (1989), Concentration and Price, Cambridge, MA: MIT Press.