DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 01-14

Market Concentration and Innovation: New Empirical Evidence on the Schumpeterian Hypothesis

Philip G. Gayle
Department of Economics, University of Colorado at Boulder
Boulder, Colorado

October 2001

Center for Economic Analysis
Department of Economics

University of Colorado at Boulder
Boulder, Colorado 80309

Market Concentration and Innovation: New Empirical Evidence on the Schumpeterian Hypothesis

Philip G. Gayle
Department of Economics
University of Colorado at Boulder
Boulder, CO 80309-0256
e-mail: philip.gayle@colorado.edu

October 2001

Abstract

This paper conducts a new empirical examination of the Schumpeterian hypothesis that more concentrated industries stimulate innovation. It is found that the lack of evidence for the hypothesized relationship in recent empirical work is largely due to the use of simple patent counts as the measure of innovative output. When citation-weighted patent count, arguably a more accurate measure of innovative output, is used, this paper finds empirical evidence in support of the Schumpeterian hypothesis.

*I would especially like to thank Prof. Bronwyn H. Hall at UC Berkeley for her expeditious provision of the data used in this research. In preparing this paper I benefited from invaluable comments from Professors Robert Mcnown, Yongmin Chen, Jose Canals-Cerda, Randy Walsh and Nicholas Flores. I am also indebted to Thitima Puttitanun for taking the time to proof read and offer valuable comments. I am responsible for any remaining errors.

1. Introduction

An important issue in economics is how market structure affects innovation. In his seminal contribution, Schumpeter (1942) claimed that society must be willing to put up with imperfectly competitive markets in order to achieve rapid technical progress. He argued that large firms in imperfectly competitive markets are the most conducive conditions for technical progress. To the extent that firms in more concentrated industries operate in a way that more closely approximates imperfectly competitive markets in which firms possess market power, this led to the long-standing and much debated hypothesis that more concentrated industries[1] are more conducive for innovation.

The Schumpeterian hypothesis challenged conventional economic thinking on the ideal market structure for optimal resource allocation and sparked a preponderance of both theoretical and empirical papers on the topic. A review of the empirical literature up to the late 70's by Kamien and Schwartz (1982) revealed an inconclusiveness of the relationship between market structure and innovative activity[2]. Results ranged from finding that imperfectly competitive markets are better at stimulating innovative activity (support for Schumpeterian hypothesis), to finding the complete opposite. Subsequently, researchers such as Geroski (1990), Blundell, Griffith and Van Reem (1995), Levin, Cohen and Mowrey (1985), and Cohen, Levin and Mowery (1987), among others, have found disproportionate evidence against the Schumpetarian hypothesis. These newer studies argued that technological opportunity, which varies across industry, is an important determinant of innovative activity and must be controlled for when investigating the relationship between market structure and innovation. They used various methods to control for these technological opportunities and point to this as the main reason that swung the evidence against the Schumpetarian hypothesis.

In this paper, I shall argue that the measure of innovative output plays a key role in testing the Schumpeterian hypothesis. The existing studies have relied heavily on the number of patents awarded (simple patent count) as a measure of innovative output[3]. Using a more accurate measure of innovative output, citation-weighted patent count, I

[1] The larger the percentage of industry output controlled by leading firms, the larger is industry concentration [see Tirole (1988), pp. 221, for measures of concentration].
[2] Cohen and Levin (1989) also provide a good review of the literature.
[3] Other measures used include number of important innovations and sales of new products.

show that the empirical evidence supports the Schumpeterian hypothesis, even after controlling for both observable and unobservable industry and firm specific characteristics which includes technological opportunity, normally cited as critical in testing Schumpeter's hypotheses.

It is suggested that monopoly power interacts with a firm's decision to innovate via *anticipated* and *current* possession of monopoly power [Kamien and Schwartz (1982)]. Innovators will have more incentive to innovate the greater the anticipated monopoly power associated with the post-innovation industry. The promised extraordinary profits in the future will more than compensate for the current R&D investment. Thus it is not controversial in the literature that greater *anticipated* monopoly power stimulates greater innovative activity. Where controversy creeps in is whether *current* possession of monopoly power stimulates greater innovation. There are theoretical arguments that posit both positive and negative relations between current monopoly power and innovative activity.

There are several arguments why the current possession of monopoly power should result in greater innovative activities. First, monopoly power with respect to current products may be extendable to new products, for example, through a dominant firm's command over channels of distribution etc. With the ability to extend monopoly power to new products, a current monopolist should find innovation more attractive. Second, as suggested by Arrow (1962), due to moral hazard problems, there may be a need to finance innovation internally, which puts firms with monopoly power at an advantage since these firms may have supernormal profits. Third, firms with current monopoly power usually have more resources and thus more likely to hire the most innovative people. Of course the third reason is related to the imperfect capital market argument underlying the second reason.

There are also disadvantages to current monopoly power in performing innovation. First, monopoly may regard additional leisure as superior to additional profits. This may be due to the lack of active competitive forces and thus generates an x-inefficiency effect. Second, a firm realizing monopoly profits on its current product or process may be slower in replacing it with a superior product or process than a newcomer. This is because the firm realizing monopoly profits on its current product calculates the profit from innovation as the difference between its current profits and the profits it could realize from the new product, whereas the newcomer regards the profits from the new product as the gain (see Kamien and Schwartz (1982)). As such, the larger current monopoly profits are, the less incentive the monopolist has to replace his own product or process.

Theoretical models comparing an incumbent's and an entrant's incentives to innovate also give mixed predictions about the impact of monopoly power on innovative effort. Gilbert and Newbery (1982) suggest that a monopolist has more incentive to win a patent race because its win avoids dissipation of rents that would occur if an entrant wins the patent race. Other theoretical models, including Reinganum(1983), Chen (2000), and Gayle(2001), suggest that factors such as uncertainty in the innovation process and the strategic relation between new and existing products may motivate entrants to spend more on R&D relative to incumbents.

Since there are forces both in favor of and against a positive relation between monopoly power and innovative activity, the net result is an empirical matter. To the extent that pure monopoly is rare in the real world, existing empirical studies have focused on the relation between market concentration and innovation, with the underlying assumption that firms in more concentrated markets tend to have more market power. The present paper will take the same approach to revisit the empirical evidence on the Shumpeterian hypothesis.

The rest of the paper is organized as follows. Section 2 discusses the measurement of innovative output. I suggest that a more precise measure of innovative output, citation-weighted patent count, can be used to test the empirical relation between market concentration and innovation. Section 3 discusses the data, section 4 presents the empirical model, section 5 discusses estimation and results, and section 6 concludes.

2. Measure of Innovative Output

For a long time now, researchers have recognized that simple patent count is not a very accurate measure of innovative output[4]. Simple patent count as a measure of innovative output has been used extensively in the empirical literature, (for review see Griliches, 1990). One reason why simple patent count is not an accurate measure of innovative output is that the technologies covered by patents are very heterogeneous in their economic and social value while, simple patent count values all patented innovations equally. Recognizing this problem, Pakes (1986), Pakes and Schankerman (1984), and Schankerman (1991), among others, attempted to measure the quality rather than quantity of innovative output using patent renewal data. In many patenting regimes, patent holders must pay an annual renewal fee in order not to forfeit the patent before its statutory limit of protection (approximately 20 years). The patent renewal literature posits that information on the value of patents can be extracted from patent renewal patterns since rational agents will only renew patents if the benefit of renewal is greater than the cost. This literature finds that a substantial number of patents were not renewed to the full statutory limit. Estimation of these models required fairly lengthy time series to observe each cohort of patents and their respective drop out dates. The majority of these models were estimated on European data rather than U.S. data, probably because U.S. only started requiring patentees to pay a renewal fee in 1982. In other words, many of the patent cohorts in U.S. data were not observable for the full statutory limit.

Recently, using U.S. data, Jaffe and Trajtenberg (1996), Hall, Jaffe and Trajtenberg (2000), and Lanjouw and Schankerman (1999) have found more creative ways to measure the value of patents by using the number of citations received by a patent. An inventor must cite all related prior U.S. patents in the application process, much like how authors of journal articles must cite related previous research. A patent examiner is responsible for insuring that all appropriate patents have been cited. Again this is analogous to the academic world where referees of journals are responsible for ensuring that all appropriate research has been cited. These citations help to define the rights of the patentee. Researchers have posited that the number of citations that a patent receives can be used to measure the relative value/importance of the technology protected

[4] See Lanjouw, Pakes and Putnam (1998).

by the patent. As such, researchers have developed new and more precise measures of innovative output using patent citations. Once more, this idea is analogous to how we measure the relative importance of published research articles. The more citations that a research article receives the more likely it is that the cited article has made an important contribution to the literature.

The measure of innovative output used in this paper is citation-weighted patent counts, that is, each patent count is weighted by the number of citations received. A brief description of the construction of the citation-weighted patent count variable is as follows. Let $n(t,\ s)$ be number of cites received at time s to patents applied for at time t. Therefore, $n(t) \ = \ \sum_{s=t}^{T} n(t,\ s)$ is the total number of cites to patent applied for at time t. Thus the time interval over which cites are counted for patents applied for at time t is T-t. The same length time interval is used to count citation information for each patent, irrespective of application date, in order to allow for comparable measures. For example, if an interval of ten years is used, then the citation measure is number of cites received by a patent within ten years after application date. The variable $n(t)$ is citation-weighted patent count. This measure of innovative output treats each patent as if it is worth the number of citations received. Thus a measure of total innovative output in a given year is the sum of citations over all the patents applied for in that year. $n(t)$ is calculated for each firm for each year in the dataset. [5]

3. Data

The dataset used in this paper is the NBER-Case-Western University R&D patents data set [see in references, Trajtenberg, Manuel, Adam Jaffe, and Bronwyn Hall (2000)]. This is a new and comprehensive dataset containing over 4800 U.S. Manufacturing firms over the period 1965 to 1995. The dataset contains usual firm specific data (2-digit industry code, sales, R&D expenditure, advertising expenditure, capital stock, assets, Tobin's q etc.) along with firms' patenting activities. Firm specific patenting information includes number of patents applied for in a given year that were

[5] For a more detailed derivation of the citation-weighted patent stock measure used in this paper see Hall, Jaffe and Trajtenberg (2000).

eventually granted and the total number of cites received by those patents. The dataset contains citation information starting only from 1976. As such, the sample used for analysis in this paper starts from 1976. Summary statistics and simple correlations of the variables used in this study are shown in tables 1 and 2. A list of broad industry categories covered by the dataset is presented in table A1. in the appendix.

Table 1

Descriptive Statistics

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
Year	-		1976	1992
Firm R&D expenditure (M$)	19.04	126.01	0	5201
Firm advertising expenditure (M$)	14.85	80.31	0	2693
Industry level R&D expenditure (M$)	2437.62	2831.99	2.08	13187.46
Firm sales (M$)	819.67	3807.70	0.001	108107
Capital stock (M$)	619.30	3026.73	0.045	95607.25
Number of Patent application	7.72	39.96	0	1303
Number of Cites to patents	37.38	217.70	0	6081
Market share	0.012	0.04	2.36e-08	0.899
Industry Concentration (Herfindahl index)	0.12	0.099	0.024	0.967

N=33250

Table 2

Correlation Matrix

	Firm R&D	Firm Sales	Patents	Cites	Industry R&D	Market Share	Industry concentration	Advertising Expenditure
Firm R&D	1							
Firm Sales	0.72	1						
Patents	0.73	0.59	1					
Cites	0.62	0.52	0.88	1				
Industry R&D	0.12	0.01	0.07	0.04	1			
Market Share	0.66	0.68	0.60	0.59	-0.09	1		
Industry Concentration	0.08	0.02	0.04	0.06	0.49	0.02	1	
Advertising Expenditure	0.36	0.41	0.34	0.27	-0.01	0.48	-0.01	1

There are some points worth mentioning about the correlations shown in table 2. First, a firm's R&D spending is positively correlated with both its simple patent count and citation-weighted patent count. In fact, these correlations are among the largest displayed in the matrix. Second, more concentrated industries, as measured by the Herfindahl index[6], are also more R&D intensive as exemplified with a correlation coefficient of 0.49. Third, industry concentration is slightly more highly correlated with firm level R&D expenditure than with innovative outputs (simple patent count and citation-weighted patent count). This suggests that industry concentration might influence innovation indirectly through R&D expenditure. Many empirical papers have

[6] The Herfindahl index is calculated by $C_t = \sum_{i=1}^{n} s_{it}^2$ where s_{it} is firm i market share at time t, and n is the number of firms in the particular industry [see Tirole (1988), pp. 221 for more on concentration indices]. Industry subscripts are suppressed for notational convenience but note that the index is calculated for each industry at each time period.

posited a direct rather than indirect relationship between industry concentration and innovation. The theoretical structure of the model in this paper posits an indirect relationship between industry concentration and innovation as suggested by the data. The last point I want to mention before moving on to the next section is that advertising expenditure is positively correlated with firm's market share as expected, but the correlation between firms market share and innovation is even higher. This seems to suggest that successful innovation could be a stronger determinant of market share compare to advertising expenditure.

4. The Empirical Model

The econometric model consists of three equations, one for research, one for innovation and one that takes account of the endogenous effect of innovation on market share. Each equation uses a different econometric treatment much like in Crepon, Duguet and Mairesse (1998). The first equation models the magnitude or intensity of research activities and is given by:

$$r_{it}^{*} = g s_{it} + b_1' x_{1it} + m_{1i} + e_{1it} \qquad (1)$$

where r_{it}^{*} is the true research intensity of firm i at time t, s_{it} is firm i market share at time t, g is the corresponding coefficient, x_{1it} is a vector of explanatory variables, b_1 the corresponding coefficient vector, m_{1i} controls for firm specific effect, and e_{1it} a random error term. In this equation the right hand side variables are firm and industry characteristics such as firm's market share, firm size, and industry concentration/competitiveness.

Having controlled for industry competitiveness and firms market share, we would expect larger firms to be more R&D intensive[7]. As such, the sign of the coefficient on firm size is expected to be positive. As stated in the introduction, more recent empirical literature swung the balance of evidence against the Schumpeterian hypothesis. That is, recent evidence suggests that industry concentration either have no effect or have a

[7] See Cohen and Klepper (1996), Scherer (1965a, 1965b)

negative impact on innovation [Geroski (1990), Blundell, Griffith and Van Reem (1995), Levin, Cohen and Mowrey (1985)]. In the structural model of this paper I have posited that industry concentration directly influences firms' R&D intensity, which in turn affects firms' level of innovation (this will be more apparent when I specify equation 2). As such, the effect of industry concentration on innovation is indirect.

Following Crepon, Duguet and Mairesse (1998), and Blundell, Griffith and Van Reem (1995), I also include firm's market share since these previous studies found that market share is a significant determinant of innovative effort. A firm's market share can also be viewed as a measure of dominance and thus theoretically should affect a firm's R&D intensity. Crepon, Duguet and Mairesse (1998) found a positive and statistically significant coefficient for the effect of market share on R&D intensity. Blundell, Griffith and Van Reem (1995) also found a positive and statistically significant coefficient but this is for the effect of market share on innovative output. As will be explained later, the structural parameters of equation 1 will not be estimated because we are more interested in the resulting parameters when equation 1 is combined with equation 2.

Equations 2 is the innovation equation and is modeled as a random-effects negative binomial regression given by:

$$E\left(n_{it} \;\middle|\; r_{it}^{*},\, x_{2it},\, m_{2i},\, e_{2it};\, a,\, b_{2}\right) \;=\; \exp\left(a r_{it}^{*} \;+\; b_{2}' x_{2it} \;+\; m_{2i} \;+\; e_{2it}\right) \qquad (2)$$

where n_{it} is citation-weighted patent count of firm i in year t. Since the dependent variable falls in the category of count data (only integer values), we specify the equation as a heterogeneous count data process conditional on research intensity and other variables. Recall that r_{it}^{*} is our R&D intensity variable from equation 1. x_{2it} is a vector of explanatory variables, m_{2i} controls for firm specific effect (heterogeneous ability to innovate), and e_{2it} is a random error term. Since x_{2it} only contains one variable which is industry level R&D, the right hand side variables in equation 2 are firm level R&D spending and industry level R&D spending. Based on previous studies such as Crepon, Duguet and Mairesse (1998), Pakes and Greliches (1984), Lanjouw and Schankerman

(1999), we expect firm's R&D spending to be positively correlated with innovation. As such, the sign of the coefficient on firm level R&D intensity should be positive.

We want to emphasize the importance of m_{2i} in equation 2. Empirical studies have found that industries vary with respect to their technological opportunities and appropriability conditions. Technological opportunities include factors such as the technological base of an industry, that is, what is the body of scientific knowledge relevant to research in an industry and how easily can this knowledge be accessed. Geroski (1990), Levin, Cohen and Mowrey (1985), Cohen, Levin and Mowery (1987), Blundell, Griffith and Van Reem (1995), stress the importance of controlling for technological opportunities and appropriability conditions when testing the Schumpeterian hypothesis. In fact, these papers show that whether you control for these factors makes the difference between accepting and rejecting the Schumpeterian hypothesis. The problem is that these factors are generally not observable. Levin, Cohen and Mowrey (1985), and Cohen, Levin and Mowery (1987) made an attempt to measure these factors via survey data. Geroski (1990) controlled for these effects via the usual fixed effect procedure applicable to panel data. Without good measures for these factors, he argued that the usual fixed or random effect procedures done with panel data are appropriate. Therefore, this explains the importance of m_{2i} in equation 2 which is also a feature of the empirical model found in Geroski (1990) and many other papers. In fact, it is a general theme in all of our equations to control for unobservable specific effects.

Following Crepon and Duguet (1997), industry level R&D is used to measure R&D externalities among firms in the same industry. According to Katz and Ordover (1990), two main types of externalities have been reported in the theoretical literature: a "competitive" externality and a "diffusion" one. Theoretical models by Loury (1979), Lee and Wilde (1980), Delbono and Denicolo (1991), Gayle (2001), all incorporated competitive externalities via a patent race, where firms invest in R&D aiming to be the first to discover an innovation. In these models, winning depends on both individual and competitors' R&D investment: a rise in a firm's R&D spending, ceteris paribus, increases its probability of winning and lowers that of its competitors. This would suggest a negative sign for the coefficient on industry level R&D in equation 2. Other theoretical models such as Katz (1986) examine diffusion externalities. In these models a firm benefits from other firms' R&D through a "spillover" effect. As such, a firm's probability of success in innovation is enhanced by more R&D of other firms in the industry. This suggests a positive sign of the coefficient on industry level R&D in equation 2. Therefore, in general theory is inconclusive as to what sign we should expect for the coefficient on industry level R&D in equation 2.

Equation 3 models the effect of innovation on market share and is given by:

$$s_{it} \;=\; \boldsymbol{j} \;+\; \boldsymbol{b}_3 n_{it} \;+\; \boldsymbol{f} a_{it} \;+\; \boldsymbol{m}_{3i} \;+\; \boldsymbol{e}_{3it} \qquad (3)$$

where s_{it} is firm i market share at time t, $\boldsymbol{j}$ is an intercept coefficient, n_{it} is citation-weighted innovation count from equation 3, a_{it} is the log of firm i advertising expenditure at time t, $\boldsymbol{m}_{3i}$ controls for firm specific effect, and $\boldsymbol{e}_{3it}$ is a random error term. Equation 3 is estimated by the usual random effects procedure when the dependent variable is continuous and normally distributed. Specification of equation 3 is a direct attempt to model the endogeneity of the relation between innovation and market structure. From equation 1 we see that a firm's market share affects it's R&D intensity which in turn influence the firm's probability of successful innovation as seen in equation 2. However, equation 3 recognizes that successful innovation in turn affects a firm's market share. We would expect that successful innovation increases a firm's market share. Also it is expected that a firm's market share should increase with its advertising expenditure since that's usually the goal of advertising. What is interesting is that we can use equation 3 to compare the relative importance of successful innovation to advertising in affecting market share.

Having specified each equation, I close this section by collecting all the equations that summarizes the full structural model as follows:

$$r_{it}^{*} \;=\; \boldsymbol{g} s_{it} \;+\; \boldsymbol{b}_1' x_{1it} \;+\; \boldsymbol{m}_{1i} \;+\; \boldsymbol{e}_{1it} \qquad (1)$$

$$E\left(n_{it} \;\middle|\; r_{it}^{*},\, x_{2it},\, \boldsymbol{m}_{2i},\, \boldsymbol{e}_{2it};\, \boldsymbol{a},\, \boldsymbol{b}_2\right) \;=\; \exp\left(\boldsymbol{a} r_{it}^{*} \;+\; \boldsymbol{b}_2' x_{2it} \;+\; \boldsymbol{m}_{2i} \;+\; \boldsymbol{e}_{2it}\right) \qquad (2)$$

$$s_{it} \;=\; \boldsymbol{j} \;+\; \boldsymbol{b}_3 n_{it} \;+\; \boldsymbol{f} a_{it} \;+\; \boldsymbol{m}_{3i} \;+\; \boldsymbol{e}_{3it} \tag{3}$$

5. Estimation and Results

Recall that the main interest of this paper is to explore how firm and industry characteristics, especially industry concentration, affects firms' innovation, where innovation can either be measured by simple patent count (standard in the literature) or citation-weighted patent count. To conduct this analysis we plug equation 1 into equation 2. This allows us to obtain an equation that expresses innovative output as a function of industry concentration among other variables. Having plugged equation 1 into equation 2, the main equation of interest is:

$$E\left(n_{it} \;\middle|\; C_t,\, s_{it},\, x_{2'it},\, \boldsymbol{m}_{2i},\, \boldsymbol{e}_{2'it};\, \boldsymbol{v},\; \boldsymbol{l},\; \boldsymbol{b}_{2'}\right) \;=\; \exp\left(\boldsymbol{v}C_t \;+\; \boldsymbol{l}s_{it} \;+\; \boldsymbol{b}'_{2'}x_{2'it} \;+\; \boldsymbol{m}_{2'i} \;+\; \boldsymbol{e}_{2'it}\right) \quad (2')$$

where C_t measures industry concentration at time t, $x_{2'it}$ is a vector of explanatory variables which includes firm size and industry level R&D. In equation 2′ the sign of $\boldsymbol{v}$ is our main interest[8]. If $\boldsymbol{v} > 0$, then there is support for the Schumpeterian hypothesis but $\boldsymbol{v} \leq 0$ is a rejection of the hypothesis. n_{it} is measured either by simple patent count or by citation-weighted patent count. The full model to be estimated consists of equations 2′ and 3. Thus there are now only two endogenous variables, s_{it} and n_{it}.

In any simultaneous equation system, two major concerns are the problem of simultaneity bias and the issue of identification. First, I discuss the problem of simultaneity bias then move on to the issue of identification.

Broadly speaking, there are two approaches to estimating the model that solves the problem of simultaneity bias. One approach involves estimating each equation separately, using a limited information estimator. Another approach is to use a full

[8] We could have gone the route of specifying both a direct and an indirect effect of market concentration on innovative output by initially including the market concentration variable in both equations 1 and 2. After plugging equation 1 into equation 2, $\boldsymbol{v}$ would then be interpreted as the total effect comprising both a direct and indirect effect. Note that the nature of the analysis would not change if this route had been chosen.

information system estimator. In both approaches we can find estimators that are consistent but, in general, full information estimation is more efficient. A full information system estimation of the model requires writing down a likelihood function for the system. As noted in Lee L.-F (1981), full maximum likelihood estimation of a simultaneous model with latent dependent variables are too complicated to be useful. To confound a full maximum likelihood estimation procedure of the model above, each equation has unobservable specific effect parameters and one of the endogenous variables is a count data variable.

Thus for practical purposes I am forced to consider a single-equation limited information approach that yields consistent estimates. The procedure used, suggested by Lee L.-F (1981), is analogous to two-stage least squares. First, the procedure requires the system to be expressed in reduced-form, that is, endogenous variables are expressed as functions of only exogenous variables. These reduced-form equations are then estimated and predicted values of the dependent variables recovered. For example, n_{it} is expressed as a function of all the exogenous variables in the model then reduced-form parameters are estimated using a random-effects negative binomial model. The reduced-form parameters are used the get predicted values of n_{it}. Predicted n_{it} is then used in the estimation of equation 3 instead of using n_{it}. Likewise, before equation $2^{/}$ is estimated we get predicted values of s_{it} from the reduced-form estimation of the s_{it} equation. Since s_{it} is a continuous variable, a normally distributed error term is assumed for the reduced-form estimation. Predicted s_{it} is then used in the estimation of $2^{/}$. Equation $2^{/}$ is estimated as a random effects negative binomial model.

Having outlined the estimation strategy, let me briefly discuss identification issues. Each of the two equations in the system includes both endogenous variables. s_{it} is on the right hand side of equation $2^{/}$ while n_{it} is on the right hand side of equation 3. Equation $2^{/}$ is identified if equation 3 has at least one exogenous variable that is not in equation $2^{/}$. The exogenous variable that identifies equation $2^{/}$ is advertising expenditure found in equation 3. Equation 3 is also identified because there are several exogenous variables in equation $2^{/}$ that is excluded from equation 3.

Following standard estimation procedures that are usually used to reject the Schumpeterian hypothesis, this paper shows that using a more precise measure of innovative output (citation-weighted patent count) can overturn previous results (i.e. find support for the Schumpeterian hypothesis). The results when innovative output is measured by simple patent count are presented in table 3 while the results when the measure is citation-weighted patent count are presented in table 4. In both tables 3and 4, the first column displays the negative binomial equation for innovation results, and the second column displays the effects of successful innovation and advertising on market share.

Table 3

Model estimates

Model	**Simple patent counts** n_{it} (1)	**Market Share** s_{it} (2)
Industry Concentration, C_t	-1.17*	-
	(0.16)	
Market share, s_{it}	13.96*	-
	(3.73)	
Firm size (log of Sales)	0.244*	-
	(0.03)	
Industry level R&D expenditure	0.00003*	-
	(5.43e-06)	
Simple patent counts, n_{it}	-	0.013*
		(0.0004)
Firm advertising expenditure (in logs)	-	0.002*
		(0.0002)
R-squared	-	0.21

Standard errors are in parenthesis.

*indicates statistical significance at the 5% level.
All regressions are fitted with a constant

Table 4

Model estimates

Model	**Citation-Weighted patent counts** n_{it} (1)	**Market Share** s_{it} (2)
Industry Concentration, C_t	1.45*	-
	(0.18)	
Market share, s_{it}	9.39*	-
	(4.04)	
Firm size (log of Sales)	0.29*	-
	(0.028)	
Industry level R&D expenditure	-0.00001*	-
	(5.97e-06)	
Citation-Weighted patent counts	-	0.02*
n_{it}		(0.0004)
Firm advertising expenditure (in	-	0.002*
logs)		(0.0002)
R-squared	-	0.25

Standard errors are in parenthesis.
*indicates statistical significance at the 5% level.
All regressions are fitted with a constant.

Column 1 of tables 3 and 4 display the main result of this paper. In column 1 of table 3 we see that the coefficient on concentration is negative. This is consistent with the newer empirical findings when innovative output is measured by simple patent count.

This is evidence against the Schumpeterian hypothesis. That is, as industries become more concentrated innovation is reduced. If we turn to column 1 of table 4 where innovative output is measured by citation-weighted patent count, then we can see that the sign of the coefficient on industry concentration switches to positive. The results in table 4 are thus consistent with the Schumpeterian hypothesis that more concentrated industry encourage innovation. It is worth emphasizing that estimation procedure and all the variables are the same in column 1 of tables 3 and 4 with the exception of the measure of innovative output.

The switch in sign on the concentration variable begs a plausible explanation. Since I argue that simple patent count is not an accurate measure of innovative output, why do we observe a significantly negative coefficient on concentration in table 3? In other words, simple patent count must be a fairly accurate measure of some process that is negatively related to industry concentration. While there might be several processes at work that drive the result, I will offer an argument that is both consistent with the data and traces back to the core of Schumpeter's argument as to why large firms in imperfectly competitive markets have an advantage in the innovative process.

One criticism of simple patent count as a measure of innovative output is that the measure captures patenting of minor technologies that can hardly be considered innovative. Significant innovations (innovations that have bigger impact), of which the citation-weighted patent count is a good measure, tend to require substantial resources that only large firms are likely to have. More concentrated industries tend to be characterized by large firms who are more able to produce these innovations. On the other hand, less concentrated industries tend to have more small firms who tend to lack the resources for major innovation, but can still produce minor innovations, and a higher number of firms tends to produce more of these minor innovations. This reasoning fits the original idea behind Schumpeter's argument why more concentration would facilitate innovation, if what he thinks is that important innovations tend to require significant resources that only large firms tend to possess.

Based on the arguments above, a negative sign on the concentration coefficient when simple patent count is used as the measure of innovative output is not surprising. The simple patent count measure is picking up a lot of minor patenting that is more prevalent in less concentrated industries. Citation-weighted patent count is designed to purge simple patent count of patents that cover minor technologies that can hardly be considered innovative. As such, citation-weighted patent count should give us a more accurate measure of the relationship between industry concentration and innovation.

It is possible to further verify that the data is consistent with these arguments. Recall that the citation-weighted patent measure is obtained by summing up citations received by a patent. Thus the citation-measure of a patent that is never cited is zero. A sufficient condition to conclude that a firm has patents that are never cited is to check if the citation-weighted patent count is less than the corresponding simple patent count. I proceed by selecting two industries that have contrasting levels of concentration from the dataset. The first industry, Motor Vehicle, is consistently among the five most concentrated industries between 1976 and 1992, and the second industry, Textile, Apparel and Footwear, has consistently been among the least concentrated over the same period. It turns out that the rate at which minor patents are applied for is almost three times (2.83 times) higher in the Textile, Apparel and Footwear industry compared to the Motor Vehicle industry[9]. This is a clear example where less concentrated industries tend to patent more minor innovations.

There are other interesting results in column 1 of tables 3 and 4. The positive sign of the coefficients on market share and firm size suggest that dominant and large firms tend to be more innovative. This finding is consistent with Blundell, Griffith and Van Reenen (1995). The sign of the coefficient on industry level R&D is positive in table 3 but changes to negative in table 4. Since table 4 has the preferred measure of innovation, I will take the negative sign on industry level R&D to be consistent with the view that "competitive" externality in R&D dominates "diffusion" type R&D externality. Recall that in section 4 we said that "competitive" externality in R&D meant that a rise in a firm's R&D spending, ceteris paribus, increases its probability of winning and lowers that of its competitors. Put another way, a rise in a firm's competitors' R&D should lower the said firm's probability of successful innovation. Crepon and Duguet (1997) in their

[9] In the Textile, Apparel and Food Industry, approximately 17% of the observations had the citation-weighted patent count measure being less than the simple patent count measure. On the other hand, in the Motor Vehicle industry only a mere 6% of the observations had citation-weighted patent count being less than the simple patent count measure.

empirical paper also found that an increase in a firm's competitors' R&D reduces its probability of successful innovation. The negative sign on industry level R&D in column 1 of table 4 supports the patent race type theoretical models found in Loury (1979), Lee and Wilde (1980), Delbono and Denicolo (1991), and Gayle (2001).

Column 2 in both tables 3 and 4 also presents some interesting results. The coefficients on innovation and advertising are both positive across both tables. Not surprisingly, this suggests that both successful innovation and advertising are strategic tools that can be used to increase market share. Since innovation and advertising expenditure are measured in different units, both coefficients must be adjusted appropriately to facilitate a meaningful comparison of relative size. The standard method[10] to adjust these coefficients is given by $\hat{b}_i^* = \frac{\hat{b}_i s_{x_i}}{s_y}$, where $\hat{b}_i^*$ is the adjusted coefficient, $\hat{b}_i$ is the unadjusted coefficient that appears in the regression, s_{x_i} is the sample standard deviation of independent variable x_i (innovation and advertising expenditure), and s_y is the sample standard deviation of the dependent variable of the regression (which in this case is market share). After applying these adjustments to the coefficients in column 2 of table 4, the adjusted coefficients on innovation and advertising expenditure are 0.296 and 0.058 respectively. In table 3 the corresponding adjusted coefficients were 0.231 and 0.070 respectively. Thus across both tables the adjusted coefficient on innovation is larger than the adjusted coefficient on advertising. This implies that, on average, successful innovation is more powerful in increasing a firm's market share compared to advertising. This should be useful strategic information for managers.

An examination of the descriptive statistics presented in table 1 suggests that a few firms account for large maximum values in variables such as, number of patent applications, number of cites to patents, market share, industry concentration, firm R&D expenditure and firm sales. One concern therefore is whether the results obtain thus far are sensitive to these outliers. In an attempt to check how robust results are to possible

---

[10] See Ramanathan (1989), "Introductory Econometrics with Applications", *Harcourt Brace Jovanovich, Inc.*, pp. 160

outliers, a total of 129 observations were deleted and the model re-estimated. Estimates based on this smaller sample are presented in tables A2. and A3. found in the appendix. All qualitative results in the full sample remain robust to the exclusion of outliers.

6. Conclusion

This paper has revisited the empirical evidence on the relationship between market concentration and innovation. It has found that a more concentrated industry stimulates innovation, in support of the Schumpeterian hypothesis. It also shows that the reason that this result has eluded recent empirical work is largely due to the use of an inaccurate measure of innovative output (simple patent count). Once innovative output is measured by citation-weighted patent count, arguably a more precise measure, a positive empirical relation between concentration and innovation is established. In addition, the empirical results support "competitive" externalities in R&D often seen in patent race models; and suggest that, on average, successful innovation is more powerful than advertising at increasing a firm's market share.

Appendix

Table A1.

2-Digit Industry code	**Industry**
01	Food & Tobacco
02	Textile, apparel & footwear
03	Lumber & wood products
04	Furniture
05	Paper & paper products
06	Printing & publishing
07	Chemical products
08	Petroleum refining & products
09	Plastics & rubber products
10	Stone, clay & glass
11	Primary metal products
12	Fabricated metal products
13	Machinery & engines
14	Computer & com. Equipment
15	Electrical machinery
16	Electronic inst. & comm. Equipment
17	Transportation equipment
18	Motor vehicle
19	Optical & medical instruments
20	Pharmaceuticals
21	Misc. manufacturing
22	Soap & toiletries
23	Auto parts

Table A2.

Model estimates

Model	**Simple patent counts** n_{it} (1)	**Market Share** s_{it} (2)
Industry Concentration, C_t	-0.76*	-
	(0.12)	
Market share, s_{it}	18.65*	-
	(4.60)	
Firm size (log of Sales)	0.23*	-
	(0.03)	
Industry level R&D expenditure	0.00003*	-
	(5.46e-06)	
Simple patent counts, n_{it}	-	0.013*
		(0.0003)
Firm advertising expenditure (in logs)	-	0.0014*
		(0.0001)
R-squared	-	0.28

Standard errors are in parenthesis.
*indicates statistical significance at the 5% level.
All regressions are fitted with a constant

Table A3.

Model estimates

Model	**Citation-Weighted patent counts** n_{it} (1)	**Market Share** s_{it} (2)
Industry Concentration, C_t	1.73*	-
	(0.14)	
Market share, s_{it}	8.22**	-
	(4.93)	
Firm size (log of Sales)	0.30*	-
	(0.030)	
Industry level R&D expenditure	-0.00001*	-
	(5.91e-06)	
Citation-Weighted patent counts n_{it}	-	0.01*
		(0.0003)
Firm advertising expenditure (in logs)	-	0.002*
		(0.0001)
R-squared	-	0.30

Standard errors are in parenthesis.
*indicates statistical significance at the 5% level.
**indicates statistical significance at the 10% level.
All regressions are fitted with a constant.

References

Arrow K., (1962), “Economic Welfare and the Allocation of Resources for Inventions”, In R. R. Nelson (ed.), *The Rate and Direction of Inventive Activity, Princeton*, N.J. ,Princeton Press.

Blundell R., Griffith R., Van Reenen J., (1995), “Dynamic Count Data Models of Technological Innovation”, *The Economic Journal*, Vol. 105, pp. 333-344.

Crepon B., Duguet E., (1997), “Research and Development, competition and innovation Pseudo-Maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity”, *Journal of Econometrics*, Vol. 79, pp. 355-378.

Crepon B., Duguet E., Mairesse J., (1998), “Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level”, *NBER working paper 6696*.

Chen Y., (2000), “Strategic Bidding by Potential Competitors: Will Monopoly Persist?”, *Journal of Industrial Economics*, Vol. XLVIII, pp. 161-175, 2000.

Cohen W., Klepper S., (1996), “A Reprise of Size and R&D,” *The Economic Journal*, Vol. 106, pp. 925-951.

Cohen W., Levin R., (1989), “Empirical Studies of Innovation and Market Structure,” *Handbook of Industrial Organization*, Vol. 2, Ch. 18.

Cohen W., Levin,R., Mowery, **D.,** (1987), “Firm Size and R&D Intensity: A Re-Examination”, *Journal of Industrial Economics*, Vol. 35, pp.543-565.

Delbono F., Denicolo V., (1991**),** “Incentives to Innovate in a Cournot Oligopoly,” *Quarterly Journal of Economics*, Vol. 106, 951-961.

Gayle P., (2001), “Market Structure and Product Innovation,” *Working paper, Economics Department at University of Colorado*.

Geroski P., (1987), “Innovation, Technology Opportunity and Market Structure,” *Oxford economic papers*.

Griliches Z., (1990), “Patent Statistics as Economic Indicators: A Survey”, Journal of Economic Literature, Vol. 28, 1661-1707.

Gilbert R. J., Newbery D. M. G., (1982), "Preemptive patenting and the persistence of monopoly", *American Economic Review,* 72:514-526.

Hall B., Jaffe A., Trajtenberg M., (2000), “Market Value and Patent Citations: A First Look”, *NBER working paper, No. 7741*.

Jaffe A., Trajtenberg M., (1996), “Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations over Time and Across Institutional and Geographic Boundaries”, *Proceedings of the National Academy of Sciences*, 93: 12671-12677.

Kamien M., Schwartz N., (1982), “Market Structure and Innovation”, *Cambridge University Press*.

Katz M., (1986), “An analysis of Cooperative research and development”, *Rand Journal of Economics*, Vol. 17, pp. 527- 543.

Katz M., **Ordover J.,** (1990), “R&D Cooperation and Competition”, Brookings Papers on Economic Activity: Microeconomics, 137-191.

Lanjouw J., Schankerman M., (1999), “The Quality of Ideas: Measuring Innovation with Multiple Indicators”, *NBER working paper* No. 7345.

Lanjouw J., Pakes A., Putnam J., (1998), “How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data”, *The Journal of Industrial Economics*, Vol. XLVI, no. 4, pp. 405-433.

Lee L.-F., (1981), “Simultaneous Equations Models with Discrete and Censored Dependent Variables”, In C. Manski and D. McFadden eds., *Structural Analysis of Discrete Data with Econometric Applications*, Boston: Massachusetts Institute of Technology Press, pp. 346-349.

Lee T., Wilde L., (1980), "Market Structure and Innovation: A Reformulation" *Quarterly Journal of Economics*, pp. 429-436.

Levin R., Cohen W., Mowery D., (1985) “R&D Appropriability, Opportunity, and Market Structure: New Evidence on Some Schumpeterian Hypotheses”, *American Economic Review*, Vol. 75, pp.20-24

Loury G., (1979) "Market structure and Innovation", *Quarterly Journal of Economics*, XCIII, pp. 395-410.

Pakes A., (1986), “Patents as Options: Some Estimates of the Value of Holding European Patent Stocks”, *Econometrica*, Vol. 54, pp. 755-784.

Pakes A., Schankerman M., (1984), “The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources”, In Griliches (ed.), *R&D, Patents and Productivity*, NBER Conference Series, The University of Chicago Press Chicago London.

Pakes A., Griliches Z., (1984), "Patents and R&D at the Firm Level: A First Look", In Griliches (ed.), *R&D, Patents and Productivity*, The University of Chicago Press Chicago London.

Ramanathan R., (1989), "Introductory Econometrics with Applications", *Harcourt Brace Jovanovich, Inc.*, pp. 160.

Reinganum J.F., (1983), "Uncertain Innovation and the Persistence of Monopoly", *American Economic Review*, 73:741-748.

Schankerman M., (1991), "How Valuable is Patent Protection? Estimates by Technology Field Using Patent Renewal Data", *NBER working paper* No. 3780.

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

Scherer F. M., (1965b), "Size of firm, oligopoly, and research: A comment", *Canadian Journal of Economics and Political Science*, 31: 256-266.

Shumpeter J. A., (1942), "Capitalism, socialism, and democracy", *New York: Harper*

Tirole J., (1988), "The Theory of Industrial Organization", *The MIT Press*, Cambridge, Massachusetts London, England, Tenth printing, 1998.

Trajtenberg M., Jaffe A., Hall B., (2000), NBER-Case-Western University Patents Data. NBER, Brandeis University, UC Berkeley, and Tel Aviv University.