Too Small to Regulate

Kaushik Basu, The World Bank and Cornell University
and
Avinash Dixit, Princeton University

Version 4.0 – November 26, 2014

Abstract

The paper argues that to achieve compliance of firms with regulations such as product quality or environmental or health standards it is better to have industries with a few large corporations than numerous small firms. We construct a model to show that limited liability constraints bind more easily in competitive industries, making it harder to impose sufficiently severe penalties and costlier to send sufficient monitors. Having large corporations allows the government effectively to delegate some of its monitoring functions to the managers of the corporation. The tradeoff between this issue and the usual argument in favor of competition is considered.

JEL Codes: L10, L51

Contact information for corresponding author (Basu):
Address: The World Bank, 1818 H Street NW, Washington, DC 20433
E-mail: kbasu@worldbank.org

Contact information for other author (Dixit):
Address: Department of Economics, Princeton University, Princeton, NJ 08544
E-mail: dixitak@princeton.edu

*We thank James Matthew Trevino for capable research assistance. Dixit thanks Nuffield College, Oxford, where part of his work was done, for its excellent academic facilities and generous hospitality.

1 Introduction

An entrenched opinion in economics, going back to at least Cournot's (1838) seminal work, holds that for an economy or a market it is better to have many small firms than a few large ones. Many firms means greater competition; that usually lowers prices closer to marginal costs, and therefore achieves greater total social surplus and greater efficiency, taking the economy closer to its Pareto frontier. This view has been prevalent, and with good reason, in industrial organization theory. Recent experience from the financial crisis of 2008 brings a new dimension to the argument. When banks and finance corporations become too big, their failure has systemic implications, inflicting collateral damage on individuals who may have nothing directly to do with those banks or corporations. Governments then feel compelled to rescue these large entities in order to minimize the collateral damage, and the anticipation of such bailout promotes reckless behavior. This phenomenon gave rise to the doctrine of Too Big to Fail (TBTF). This has recently led the International Monetary Fund (2014) to propose, as one possible solution, limiting the size of banks.

The aim of this paper is to present and evaluate a counterargument, namely, that there are situations where it is better to have few large firms rather than many small ones. In itself, this is not new. There are special situations where arguments in favor of monopoly have been made. It is, for instance, believed that when it comes to creating money, it is best to have only one agent doing this in one economy, namely its central bank. And in industrial organization, complementary goods are better sold by a monopolist than by separate firms, so as to avoid double marginalization.

Here we develop a different argument, to do with the scope for and efficacy of government regulation and the agency of the state. In essence, we argue that having many firms may make them Too Small to Regulate (TSTR); conversely, having few firms makes it easier to regulate and administer them.

Consider an environmental regulation. Suppose all firms in an industry are required to install some equipment meant to prevent them from polluting the environment. To make sure that they comply, government has to send out inspectors to make spot checks. The same problem crops up in other areas. Government often has to impose minimal health and safety standards on restaurants. Just asking restaurants to do so is unlikely to be enough. It has to put together a system for monitoring and for punishing offenders. Tax authorities must similarly monitor firms and punish tax-evaders. All these problems give rise to challenges in administration.

These problems of administration and regulation are typically more acute in industries where there are numerous separate firms. In most countries, cars are manufactured by a few large firms and sold by their dealerships, and a buyer can be quite confident that the cars will meet with the standards required of them—ranging from pollution control to safety. However, second-hand car markets are characterized by thousands of small firms, so there is much greater uncertainty about these standards. In olden days, when milk was supplied to households by hundreds and thousands of small dairy 'firms,' adulteration in the form of dilution with water was rampant. A study of milk supply by small farmers and dairy producers in rural Bangladesh revealed that 100% of the milk samples were so adulterated (Chanda et al, 2012). Now, in the segment of the industry which is catered to by a few large firms, such regulatory violations are few and far between.

Examples of this kind can be found in many countries and in many sectors. Martin Wolf (2014) recently conjectured that "banking systems dominated by a few large institutions might be more stable than competitive ones." Indeed, the traditional moneylender system violated usury laws with much greater abundance than large modern banks do.

The argument that we propose in order to explain this stylized fact is simple. Consider the hamburger industry. Suppose first that there are thousands of small burger shops, each one an independent firm. There is some regulation, R, about the ingredients of burgers that they are supposed to comply with. When one of these small firms is caught violating R, there is a limit to how much such a firm can be punished. The government can for instance, take away all its profits and may be even the owner's car, but beyond a point there may be a limited liability constraint. Even without that, there are natural limits to the size of punishment on a small player. The limited threat of punishment is often inadequate to ensure compliance, using a standard crime-and-punishment argument (Becker 1968). Contrast the case where all those burger outlets are owned by one firm, MacDuck. If any of the outlets is caught violating R, the government can impose a much larger punishment on this large firm that has the capacity to pay the bigger penalty (in the extreme case, the entire profit from all its outlets). In other words, the limited liability constraint on punishments is much more binding on small firms than big firms.

However, the simple verbal statement is not enough. For example, we know from Bernheim and Whinston (1990) that in a repeated interaction with multiple contacts, good behavior is not always sustained by the threat of punishing one transgression by ending all interactions: the optimal response to such a strategy may instead be to transgress in all dimensions at the same time. Therefore it is important to prove that the drastic punishment strategy does work in the present context. We do this in the formal model of the next section, proving that it is easier for government to regulate a monopolistic industry than a competitive one.

Our result does not constitute an unequivocal case for large firms, because we have to weigh on the other side the usual costs of monopoly. But it is an important enough argument to deserve attention in the real world of policymaking. Regulation is a part of modern industry; and so creating an environment that makes it incentive compatible for firms to abide by the regulation is important for running a successful and efficient economy. In Section 4 we examine how the cost of regulation can be traded off against the usual case for greater competition.

2 Related literature

Biglaiser and Friedman (1994) consider a related situation where a middleman handling several products can be a better guarantor of quality than separate individual firms each selling one product directly to consumers, because consumers' experience of low quality in any one product purchased from the middleman causes loss of the middleman's reputation affecting all products. In our model the prospect of penalties creates the incentive, not the prospect of losing the price premium for a high-quality product. Thus we make an interesting connection between public policy and industry structure.

Dharmapala, Slemrod and Wilson (2011) consider optimal taxation of firms when the government has to incur a per-firm fixed administrative cost of tax collection. This makes it optimal to exempt a set of smallest firms from taxation, and gives rise to social costs in addition to the lost revenue, because some firms above the exemption threshold deliberately choose to stay smaller to avoid the tax. They do not pursue the implications for industrial organization. But they do conclude that "models of government policy toward small firms should address all aspects of taxation and regulation simultaneously," and our paper contributes to the regulation aspect.

3 The Model

Consider n small firms subject to some regulation concerning quality, environmental damage, etc. that is costly to comply with. Each firm's profit if it complies is π_l. If it violates the requirement and gets away with it, the profit is $\pi_h > \pi_l$.

The government has m inspectors to enforce the regulation, where $0 < m < n$. Each visits a randomly chosen firm but without any duplication (sampling without replacement). Focus on any one firm. The probability that it is visited is obviously m/n; therefore the probability that it is not visited is

$$P_1 = 1 - \frac{m}{n} = \frac{n-m}{n}$$

Suppose this firm violates the regulation. If it is not visited, it enjoys π_h; if it is visited, it suffers the maximum possible penalty. We normalize its payoff in this situation at 0.[1] Therefore its expected profit from violation is

$$P_1 \; \pi_h + (1 - P_1) \; 0 = \frac{n-m}{n} \; \pi_h$$

Successful deterrence requires this to be $< \pi_l$, or

$$(n-m) \; \pi_h < n \; \pi_l, \quad \text{or} \quad m \; \pi_h > n \; (\pi_h - \pi_l)$$

or

$$\frac{m}{n} > \frac{\pi_h - \pi_l}{\pi_h} \tag{1}$$

Now suppose there is just one large firm that owns all n branches or outlets. It can choose any subset of them, say k, to violate and the other $(n-k)$ to comply. Let P_k denote the probability that the firm is not caught.

$$\begin{aligned}
P_k \;=\; & \Pr(\text{first inspector goes to one of the complying } (n-k) \text{ branches}) \\
& \cdot \Pr(\text{second inspector goes to one of the complying } (n-k-1) \\
& \qquad \text{out of the remaining } (n-1) \text{ branches}) \ldots \\
& \cdot \Pr(m^{\text{th}} \text{ inspector goes to one of the complying } (n-(m-1)-k) \\
& \qquad \text{out of the remaining } (n-(m-1)) \text{ branches}) \\
=\; & \frac{n-k}{n} \cdot \frac{n-1-k}{n-1} \ldots \frac{n-(m-1)-k}{n-(m-1)}
\end{aligned}$$

[1] Expected utility being cardinal, the theory works just as well with any other level. The important assumption is that the large firm's punishment is exactly n times that of each small firm. If the large firm can be subjected to even greater punishment, for example by fining away the profits of a conglomerate in other lines of business, our argument will be further strengthened.

Replacing k by $(k-1)$ in this,

$$\begin{aligned} P_{k-1} &= \frac{n-(k-1)}{n} \cdot \frac{n-1-(k-1)}{n-1} \ldots \frac{n-(m-1)-(k-1)}{n-(m-1)} \\ &= \frac{n+1-k}{n} \cdot \frac{n-k}{n-1} \ldots \frac{n-(m-1)-k+1}{n-(m-1)} \end{aligned}$$

Dividing,

$$\frac{P_k}{P_{k-1}} = \frac{n-(m-1)-k}{n+1-k} = \frac{n+1-m-k}{n+1-k} \tag{2}$$

This recursion equation can be solved to evaluate P_k, starting with $P_0 = 1$. For example, if $n = 100$ and $m = 10$,

$$\begin{aligned} P_1 &= P_0 \, \frac{91-1}{101-1} = 0.9, \\ P_2 &= P_1 \, \frac{91-2}{101-2} = 0.9 \, * \, 0.898989, \\ P_3 &= P_2 \, \frac{91-3}{101-3} = 0.9 \, * \, 0.898989 \, * \, 0.897959 \ldots \\ P_{91} &= P_{90} \, \frac{91-91}{101-91} = 0 \end{aligned}$$

Consider the large firm's strategy of having k of its units violate, for $k = 1, 2, \ldots n$. If even one violation is detected, it will be fined the maximum possible amount, namely profit from all its branches, and end up with 0. Therefore the expected payoff from its strategy is

$$\begin{aligned} V_k &= P_k \, [\, k\,\pi_h + (n-k)\,\pi_l \,] + (1-P_k)\; 0 \\ &= P_k \, [\, n\,\pi_l + k\,(\pi_h - \pi_l) \,] \end{aligned}$$

Writing the corresponding expression for V_{k-1}, subtracting, and using (2), we find

$$\frac{V_{k-1} - V_k}{P_{k-1}} = \frac{m\,n\,\pi_l + [\,(m+1)\,k - (n+1)\,]\,(\pi_h - \pi_l)}{n+1-k}$$

The numerator is an increasing function of k, and the denominator is positive and a decreasing function of k, for $k = 1, 2, \ldots n$, so the ratio is an increasing function of k. Therefore, if $V_0 - V_1 > 0$, then $V_1 - V_2 > 0$ also, and by induction $V_{k-1} > V_k$ for all k. Therefore to ensure deterrence it suffices to check $V_0 > V_1$, that is

$$n\,\pi_l > P_1 \, [\, n\,\pi_l + (\pi_h - \pi_l) \,] = \frac{n-m}{n} \, [\, n\,\pi_l + (\pi_h - \pi_l) \,]$$

The condition simplifies to

$$\frac{m}{n} > \frac{\pi_h - \pi_l}{n\,\pi_l + (\pi_h - \pi_l)} \tag{3}$$

Comparing (1) and (3), the expression in the denominator on the right hand side of (3) is larger than that in (1). Therefore the minimum ratio of m/n required for deterrence of one large firm is lower than that needed with separate firms. Thus deterring one large firm from engaging in any violations requires less inspection than deterring all of separate small firms.

As an example, suppose $n = 100$ and $\pi_h = 2\,\pi_l$. From (1) the ratio of inspectors to firms with separate firms must be at least $1/2$, so the government needs at least 50 inspectors to deter all firms. From (3) the ratio with one firm needs to exceed only $1/101$, so just one inspector suffices to achieve total deterrence!

In reality a sole proprietor may have to pay a fine in excess of his/her firm's profit, and courts may judge it excessive punishment to extract all the profit of a large multi-outlet firm for a violation at just one. But qualitatively the idea that a large firm can be induced to comply with regulations at much smaller administrative cost than a large number of small firms should be quite robust.

4 Balancing Costs of Oligopoly and Regulation

Suppose there are n firms, each with f franchises or outlets. Thus the m inspectors must cover a total number $N = n\,f$ of establishments. If any of the f outlets of one firm is caught in violation of the regulation, it pays a fine equal to its profit from all f of its outlets. Then it is easy to see that the condition (3) for deterrence changes to

$$\frac{m}{N} > \frac{\pi_h - \pi_l}{f\,\pi_l + (\pi_h - \pi_l)} \tag{4}$$

Let us see how this works with the illustrative numbers we had used before. Remember that with $\pi_h = 2\,\pi_l$ when there were $n = 100$ firms each operating $f = 1$ outlet, 50 inspectors were needed to achieve deterrence. Suppose the government has only enough resources to provide $m = 6$ inspectors. Keeping $n\,f = 100$, (4) becomes

$$\frac{6}{100} > \frac{1}{f+1}, \qquad \text{or} \qquad f > 15.67\,.$$

Competition will be best served if the number of firms is as large as possible, so f should be as small as possible. To satisfy integer constraints, this requires $f = 20$ and $n = 5$. Thus 6 inspectors can achieve deterrence from 5 firms each of which operates 20 outlets.

More generally, the analysis can be extended to find an optimal tradeoff between regulation and competition. To illustrate this, assume a quadratic quasilinear utility function

$$U = x_0 + a \sum_{i=1}^{N} x_i - \tfrac{1}{2}\, b \left\{ \sum_{i=1}^{N} (x_i)^2 + \theta \sum_{i,j=1}^{N} \sum_{j \neq i} x_i\, x_j \right\} \tag{5}$$

where x_0 is the quantity of the outside (numeraire) good, and x_i for $i = 1\ 2, \ldots N$ are the quantities of the products of the industry being regulated. Also $0 < \theta < 1$. In the extreme case $\theta = 0$ the products have independent demands; in the extreme case $\theta = 1$ all products are perfect substitutes. In the intermediate range, any two products are mutual substitutes, but less than perfect substitutes, for example hamburgers sold at different locations. Then the inverse demand function for product 1 is

$$p_1 = \frac{\partial U}{\partial x_1} = a - b\, x_1 - \theta\, b \sum_{i=2}^{N} x_i$$

and similarly for the other products. Let c be the marginal cost of producing any product while conforming to the quality regulation.

Suppose firm 1 owns products 1 to f, and chooses their quantities in Cournot-Nash manner taking the quantities $x_{f+1}, \ldots x_N$ as given, to maximize its total profit

$$\Pi_{\text{firm 1}} = \sum_{i=1}^{f} (p_i - c)\, x_i$$

We have

$$\begin{aligned}
\frac{\partial \Pi_{\text{firm 1}}}{\partial x_1} &= p_1 - c + \sum_{i=1}^{f} x_i \frac{\partial p_i}{\partial x_1} \\
&= a - b\, x_1 - \theta\, b \sum_{i=2}^{N} x_i - c - b\, x_1 - b \sum_{i=2}^{f} \theta\, x_i \\
&= a - c - 2\, b \left[x_1 + \theta \sum_{i=2}^{f} x_i \right] - \theta\, b \sum_{i=f+1}^{N} x_i
\end{aligned}$$

Thus the firm internalizes the effect of the quantity of one of its products on its other products but not on the products of other firms. Setting this equal to zero and letting x denote the quantity of each product in the fully symmetric Cournot-Nash equilibrium, we have

$$\begin{aligned} 0 &= a - c - \{\, 2\,[1 + \theta(f-1)] + \theta\,(N-f)\,\}\; b\; x \\ &= a - c - b\,\{\, 2 + \theta\,(N + f - 2)\,\}\; x \end{aligned}$$

Therefore the solution is

$$x = \frac{a-c}{b\,[2 + \theta\,(N + f - 2)]} \tag{6}$$

and then the price of each product is

$$\begin{aligned} p &= a - [1 + \theta\,(N-1)]\; b\; x \\ &= a\,\left[1 - \frac{1 + \theta\,(N-1)}{2 + \theta\,(N+f-2)}\right] + c\,\frac{1 + \theta\,(N-1)}{2 + \theta\,(N+f-2)} \\ &= a\,\frac{1 + \theta\,(f-1)}{2 + \theta\,(N+f-2)} + c\,\frac{1 + \theta\,(N-1)}{2 + \theta\,(N+f-2)} \end{aligned} \tag{7}$$

The profit for each product when the firm abides by the regulation is given by

$$\pi_l = (p-c)\,x = \frac{(a-c)^2}{b}\;\frac{1 + \theta\,(f-1)}{[2 + \theta\,(N+f-2)\,]^2} \tag{8}$$

Suppose that by violating the quality regulation the firm can save s of the marginal cost of each product. We assume that the firm will go on producing the same quantity x given by (6) of each product, because increasing the quantity to optimize profit for the new lower marginal cost $c - s$ would be a giveaway of its violation. Therefore the increase in the profit from each product is

$$\pi_h - \pi_l = s\,x = \frac{s\,(A-c)}{b\,[2 + \theta\,(N+f-2)]} \tag{9}$$

The minimal number of inspectors m needed to deter violation is given by (4). Suppose w is the cost of each inspector. Using (8) and (9), the cost of regulation is

$$\begin{aligned} w\,m &= \frac{w\,N\,(\pi_h - \pi_l)}{f\,\pi_l + (\pi_h - \pi_l)} \\ &= \frac{w\,N}{1 + f\,\dfrac{\pi_l}{\pi_h - \pi_l}} \\ &= \frac{wN}{1 + f\,\dfrac{(p-c)x}{sx}} \\ &= \frac{w\,N}{1 + \dfrac{A-c}{s}\,\dfrac{f\,[1 + \theta(f-1)]}{2 + \theta(N+f-2)}} \end{aligned} \tag{10}$$

The social welfare, i.e. the utility generated by the quantities of these products net of production and regulation costs, is found by substituting from all these expressions into (5). Leaving out an inessential constant, the result is

$$
\begin{aligned}
U &= N\,A\,x - \tfrac{1}{2}\,b\,\left[\,N\,x^2 + \theta\,N(N-1)\,x^2\,\right] - N\,c\,x - w\,m \\
&= \frac{N\,(A-c)^2}{b\,[2+\theta(N+f-2)]} - \frac{1}{2}\,\frac{(A-c)^2}{b}\,\frac{N\,[1+\theta(N-1)]}{[2+\theta(N+f-2)]^2} - \frac{w\,N}{1+\frac{A-c}{s}\,\frac{f\,[1+\theta(f-1)]}{2+\theta(N+f-2)}} \\
&= N\,\frac{(A-c)^2}{b}\,\frac{2+\theta(N+f-2)-\frac{1}{2}\,[1+\theta(N-1)]}{[2+\theta(N+f-2)]^2} - \frac{w\,N}{1+\frac{A-c}{s}\,\frac{f\,[1+\theta(f-1)]}{2+\theta(N+f-2)}} \\
&= N\,\frac{(A-c)^2}{b}\,\frac{3+\theta(N+2f-3)}{2\,[2+\theta(N+f-2)]^2} - \frac{w\,N}{1+\dfrac{A-c}{s}\,\dfrac{f\,[1+\theta(f-1)]}{2+\theta(N+f-2)}} \qquad (11)
\end{aligned}
$$

Now f can be chosen from among the divisors of N to maximize social welfare.

The most interesting application is where N is large. Then we can legitimately ignore the integer constraint and treat $f/N = \phi$ say as a continuous variable. Note that ϕ can range from 0 to 1. Very small values of ϕ correspond to a large number of small firms, and $\phi = 1$ corresponds to a very large monopoly firm for all N products. In fact $1/\phi = N/f = n$, the number of firms.

Substituting into (11) and retaining the leading terms in each combination,[2] we have

$$
\begin{aligned}
U &= N\,\frac{(A-c)^2}{b}\,\frac{\theta\,N\,(1+2\,\phi)}{2\,\theta^2\,N^2\,(1+\phi)^2} - \frac{w\,N}{\frac{A-c}{s}\,\frac{\theta\,(\phi\,N)^2}{\theta\,N\,(1+\phi)}} \\
&= \frac{(A-c)^2}{2\,\theta\,b}\,\frac{1+2\,\phi}{(1+\phi)^2} - \frac{s\,w}{A-c}\,\frac{1+\phi}{\phi^2} \qquad (12) \\
&= \frac{s\,w}{A-c}\,\left[\,\frac{(A-c)^3}{2\,\theta\,b\,s\,w}\,\frac{1+2\,\phi}{(1+\phi)^2} - \frac{1+\phi}{\phi^2}\,\right]
\end{aligned}
$$

Define

$$K = \frac{(a-c)^3}{2\,\theta\,b\,s\,w} \qquad (13)$$

and

$$g(\phi) = K\,\frac{1+2\,\phi}{(1+\phi)^2} - \frac{1+\phi}{\phi^2} \qquad (14)$$

[2]For this procedure to give a good approximation, we must have $\theta\,N >> 1$. It should not be used when θ is small. Of course when $\theta \approx 0$, the products have independent demands and even a small firm enjoys a monopoly in its own market, so traditional arguments for competition are irrelevant and our argument about regulation costs prevails.

Maximizing U is equivalent to maximizing $g(\phi)$.

Now

$$\begin{aligned} g'(\phi) &= K \, \frac{(1+\phi)^2 * 2 - (1+2\phi) * 2(1+\phi)}{(1+\phi)^4} - \frac{\phi^2 * 1 - (1+\phi) * 2\phi}{\phi^4} \\ &= 2\,K \, \frac{1+\phi-(1+2\phi)}{(1+\phi)^3} - \frac{\phi - 2\,(1+\phi)}{\phi^3} \\ &= -\,2\,K \, \frac{\phi}{(1+\phi)^3} + \frac{2+\phi}{\phi^3} \\ &= \frac{1}{\phi^2} \left[\frac{2}{\phi} + 1 - 2\,K \left(\frac{\phi}{1+\phi} \right)^3 \right] \qquad (15) \end{aligned}$$

The first two terms in the square brackets constitute a decreasing function of ϕ, going from ∞ to 3 as ϕ goes from 0 to 1. The last term is an increasing function of ϕ, and goes from 0 to $K/4$ as ϕ goes from 0 to 1. Therefore we have two cases:

(i) If $K < 12$, the right hand side of (15) is positive throughout its range, so $\phi = 1$ (a monopoly) is optimal.

(ii) If $K > 12$, there is a unique ϕ^* such that $g'(\phi) > 0$ when $\phi < \phi^*$ and < 0 when $\phi > \phi^*$. Therefore ϕ^* is optimal. A smaller K shifts the function $g'(\phi)$ up and therefore implies a larger optimal ϕ^*. So when s (the firms' cost saving by violating the regulation) and w (the cost of inspection) are larger, K is smaller and larger multi-product firms are optimal. The same is true when θ is larger (the products are closer substitutes), but the intuition for this is not clear.

To get a better understanding of the effect of K in the range > 12, recall that $1/\phi = N/f = n$ is the number of firms, and write the first-order condition $g'(\phi) = 0$ as

$$\begin{aligned} K &= \tfrac{1}{2} \left(\frac{2}{\phi} + 1 \right) \left(\frac{1+\phi}{\phi} \right)^3 \\ &= \left(\frac{1}{\phi} + \tfrac{1}{2} \right) \left(\frac{1}{\phi} + 1 \right)^3 \\ &= \left(n + \tfrac{1}{2} \right) \left(n + 1 \right)^3 \qquad (16) \end{aligned}$$

Table 1 shows this function for several values of n. We see that K increases very rapidly – somewhat faster than n^3 over most of this range, and close to n^4 for large n. Actually K is the exogenous variable, so the table should be read as showing for what values of K the successive n's are optimal. As K increases, n increases much more slowly (slower than $K^{1/3}$, and like $K^{1/4}$ for large K).

Table 1: Tabulation of the function (16)

n	K
1	12.0
2	67.5
3	224.0
4	562.5
5	1188.0
6	2229.5
7	3840.0
8	6196.5
9	9500.0
10	13975.5
...	...
15	63488.0
16	81064.5
20	189850.5
25	448188.0
30	908625.5

To get better intuition about the magnitude of the key composite parameter K, decompose it as

$$K = \frac{1}{\theta}\,\frac{1}{w}\,\frac{(a-c)^2}{b}\,\frac{(a-c)/2}{s} \tag{17}$$

Each of the factors in this can be given a more intuitive interpretation.

Take the case of perfect substitutes ($\theta = 1$), where oligopoly is best able to avoid competitive price-cutting, so it is the case where the usual antitrust argument has most validity. If the industry is a monopoly, with $f = N$ or $\phi = 1$, each outlet's profit is found from (8) to be

$$\pi_l(\text{monopoly}) = \frac{(a-c)^2}{4\,b\,N}$$

The price-cost margin under monopoly is found from (7) to be

$$p - c = \frac{a-c}{2}$$

so the last factor on the right hand side of (17) is the reciprocal of the fraction of this margin that violating the quality regulation would save for the firm.

Now we can think of plausible numerical magnitudes, with the example of a hamburger chain in mind. Suppose $N = 1000$. Suppose the profit π_l of each outlet when the whole is run as a monopoly is \$500,000; this yields

$$\frac{(a-c)^2}{b} = 5 \, * \, 10^5 \, * \, 4 \, * \, 10^3 = 2 \, * \, 10^9$$

Suppose the cost of each inspector is \$100,000; this includes a reasonable efficiency wage to deter corruption, and other supporting apparatus of equipment etc. And suppose $p = 3$, $c = 2$ (this means each outlet sells 500000/365, or about 1400 burgers each day) and $s = 0.25$ (the firm can save 25 cents on each burger by cheating on the quality regulation). Substituting all these numbers in (17) yields $K = 80,000$. Then Table 1 shows that it is optimal to have 16 firms, each operating about 62 outlets. This is more competitive than the oligopoly of a few large firms and a fringe of very small firms, but far from any perfect competition of small individual stores.

In the above illustration, if we change N from 1000 to 10000, K changes to 800,000, and the table gives $n \approx 29$, making it optimal to have a few more but much larger firms, each with $10000/29 \approx 345$ outlets.

4.1 Effect of substitution

However, one puzzle remains, and its resolution leads to an interesting modification of our usual intuition about the interaction between substitution and competition. From (17) we see that the greater the substitutability between the products (higher θ), the smaller is K, and therefore n is also smaller. Thus better substitutes justify more monopoly, which runs counter to the usual intuition. The answer is to be found in a somewhat subtle interaction.

In the expression (12) for the net social welfare, the first term is the conventional oligopoly equilibrium welfare, and the second is the cost of regulation. Focus on the first, say U_O:

$$U_O = \frac{(A-c)^2}{2\,\theta\,b} \; \frac{1+2\,\phi}{(1+\phi)^2}$$

We have

$$\frac{\partial U_O}{\partial \theta} = -\,\frac{(A-c)^2}{2\,\theta\,b^2} \; \frac{1+2\,\phi}{(1+\phi)^2} < 0$$

so for a given level of concentration (as measured by ϕ), greater substitution does lower welfare, in conformity with the usual intuition. And

$$\frac{\partial U_O}{\partial \phi} \;=\; \frac{(A-c)^2}{2\,\theta\,b} \; \frac{(1+\phi)^2 \, * \, 2 - (1+2\phi) \, * \, 2(1+\phi)}{(1+\phi)^4}$$

$$= \quad - \frac{(A-c)^2}{\theta\, b} \, \frac{\phi}{(1+\phi)^3} < 0$$

so for a given level of substitution, greater concentration lowers welfare, again as expected. But the cross-effect is positive:

$$\frac{\partial}{\partial \theta}\left(\frac{\partial U_O}{\partial \phi}\right) = \frac{(A-c)^2}{\theta^2\, b} \, \frac{\phi}{(1+\phi)^3} > 0$$

Greater substitution makes the *marginal* welfare effect of increased concentration less negative. And in our combined analysis it is this marginal effect that is traded off against the marginal effect of greater concentration on reducing regulation costs. That is why, with greater substitution, it is optimal to allow greater concentration.

5 Possible Extensions

We argued that compliance with regulations such as those pertaining to product quality or environmental standards may be easier to achieve in an industry run by larger firms. In effect, the government can use the internal governance and incentive mechanisms of large corporations to undertake the tasks of enforcement of the regulations in a more cost-effective way than can the government's own monitoring system. Thus the model blurs the public-private dividing line in an interesting way.

We did not explicitly model the mechanism a firm uses to control behavior within its organization, for instance, when a large firm orders some of its outlets or branches to behave in certain ways. A large literature starting with Coase (1937) and developed by Williamson (2002) and others has established that the rationale for existence of firms is that they have ways to manage their internal decisions and behavior more efficiently than external incentives or markets. So long as the large firm's cost of internal management is less than that of government inspection, our qualitative argument will remain valid. It would be easy to add this feature and cost comparison explicitly to the model, but that will merely complicate the algebra without adding insight. What is central to our paper is limited liability, be it legally specified or naturally ordained because a small firm has nothing to offer beyond a small amount.

References

Becker, G. 1968. "Crime and punishment: An economics approach." *Journal of Political Economy,* vol. 76, no. 2, Mar.-Apr., pp. 169–217.

Bernheim, B. Douglas and Whinston, Michael. 1990. "Multimarket contact and collusive behavior." *Rand Journal of Economics,* vol. 21, no. 1, Spring, pp. 1–26.

Biglaiser, Gary and James W. Friedman. 1994. "Middlemen as guarantors of quality." *International Journal of Industrial Organization,* vol. 12, no. 4, December, pp. 509-531.

Chanda, T., Debnath, G.K., Hossain, M.E., Islam, M.A. and Begum, M. K. 2012. "Adulteration of raw milk in the rural areas of Barisal district of Bangladesh." *Bangladesh Journal of Animal Science,* vol. 41, no. 2, pp. 112–115.

Coase, Ronald H. 1937. "The nature of the firm." *Economica,* vol. 4, issue 16, November, pp. 386-405.

Cournot, Antoine-Augustin. 1838. *Recherches sur les Principes Mathématiques de la Théorie des Richesses.* Paris: L. Hachette.

Dharmapala, Dhammika, Joel Slemrod and John Douglas Wilson. 2011. "Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs." *Journal of Public Economics,* vol. 95, nos. 9–10, October, pp. 1036–1047.

International Monetary Fund. 2014. *Global Financial Stability Report: Moving from Liquidity to Growth Driven Markets.* IMF: Washington, D.C.

Williamson, Oliver E. 2002. "The theory of the firm as governance structure: From choice to contract." *Journal of Economic Perspectives,* vol. 16, Spring, pp. 171-95.

Wolf, Martin. 2014. " 'Too big to fail' is too big to ignore." *Financial Times,* April 16, p. 7.