How costly are restrictions on size?

Nezih Guner [*], Gustavo Ventura, Xu Yi

*Department of Economics, The Pennsylvania State University, Kern Graduate Building,
University Park, PA 16802, USA*

Abstract

We develop a simple framework to address government policies that restrict the size of establishments in a particular sector. The economy we study is a two-sector extension of the span-of-control model of Lucas [Lucas, R.E., 1978. On the size distribution of business firms. Bell Journal 9, 508–523]. In the model, production requires a managerial input, and individuals sort themselves into managers and workers. Since managers are heterogeneous in terms of their ability, establishments of different sizes coexist in equilibrium in each sector. We then study government policies that aim to change the size distribution of establishments in a given sector, such as Japan's Large Scale Retail Location Law. How costly are these policies? What is their impact on productivity, the number and size distribution of establishments? We find that these effects are potentially large.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

Policies that affect establishments and firms of different sizes differently are pervasive at the cross-country level. Different countries either restrict the size of large production units, or subsidize small production ones, or both. In some developing countries, such policies take extreme forms. In India, for instance, in the late 1980s, about 800 product groups were reserved for small scale firms, i.e. these goods could not be produced by large firms.

-----

[*] Corresponding author.
E-mail address: nguner@psu.edu (N. Guner).

Furthermore, these reserved product lines constituted about half of all product lines in some sectors, such as light engineering—see Little et al. (1987) for a detailed description of such policies in India. A perhaps more common practice, however, is differential tax treatment of firms of different sizes, as governments often find taxing larger firms an easier task than taxing small enterprises—see Gollin (1995) for a study of such differential tax treatment in Ghana.

These policies are by no means restricted to developing countries. Several rich countries, for example, have policies that regulate the size and operation of establishments in the retail sector. While South Korea, the UK, France and Japan have explicit restrictions on large establishments in retailing, Germany has regulations on location and operating hours in this sector—see Baily (1993) and Baily and Solow (2001). Japan is, in particular, concurrently unique among developed countries as it regulates heavily and at the national level the *size* of retail shops. These regulations have been in place for long time, and are still rather strict despite recent reforms. Given their prominence, we describe these regulations in detail below.

There is a number of observations that make the study of these policies of special relevance. First, several authors document large differences in Total Factor Productivity (TFP) across countries. It is natural to surmise that policy distortions contribute a great deal to measured differences. Nicoletti and Scarpetta (2003) find that product market regulation (measured by indicators, such as the extent of state control, barriers to entrepreneurship, barriers to trade and investment, etc.) and Total Factor Productivity growth are strongly and negatively correlated for a set of OECD countries. Recent work has built models that link policy distortions to TFP differences; Schmitz (2001); Bergoeing et al. (2002); Parente and Prescott (2000) and Restuccia and Rogerson (2003) are examples.

Second, the fact that the retail sector is distorted in several countries might be of importance. There is evidence of substantial productivity growth in the service sector, and in the retail sector in particular. According to Basu et al. (2003), the productivity growth in wholesale and retail trade between 1995 and 2000 was the second highest among all sectors in the US, second only to information technology producing sectors. Third, policies that affect the *size* of establishments are likely to be costly as large establishments account for a disproportionate fraction of output and employment. This is generally the case across different sectors, and true in special for the retail sector. Some figures are stark: we calculate from the 1997 US Economic Census that retail establishments with 100 workers or more, constituted 2.4 percent of the total number of establishments in the sector, but accounted for about 32 percent of total employment.

Finally, as is discussed in a number of studies, the regulated retail sector in Japan is special in a number of ways. First, the number of stores per-person is rather high. Flath (2003) reports that there are about 11.2 stores per 1000 population in Japan, while the same number is 6.1 in the US. Second, small retail establishments in Japan contribute disproportionately to employment in the retail sector. According to McKinsey-Global-Institute (2000), share of traditional mom-and-pop stores in total hours worked in retailing is about 55 percent in Japan and 19 percent in the US.[1] Finally, these studies point out that

---

[1] More generally, the whole size distribution of the retail sector in Japan differs dramatically from the one in the US. For instance, we calculate from Japanese census data that establishments with 100 employees or more constituted less than 1 percent of total establishments, and accounted for 12 percent of total employment in the sector in 2001.

productivity in the sector is smaller than in the United States and other sectors of the Japanese economy. McKinsey-Global-Institute (2000) documents that output per-worker in merchandise retailing in Japan was about half of the level in the US in 2000 at common prices. In comparison, aggregate output per-worker in Japan was about 70 percent of the US in 2000.

In this paper, we develop a simple framework to evaluate restrictions to the size of establishments at the sectorial level. The model economy is a two-sector generalization of Lucas (1978). There is a single representative household in the economy, which is inhabited by individuals that are heterogenous in terms of their endowment of sector-specific managerial skills. As a result of the underlying heterogeneity, individuals sort themselves between managers and workers in each sector. Furthermore, since those who become managers are heterogeneous in terms of their skills, establishments of different sizes coexist in equilibrium in each sector. We parameterize and calibrate the model to reproduce observations of the United States, which we take as a relatively distortion-free economy for the purposes of this paper. We subsequently introduce distortions on the size of establishments in one of the sectors, which we calibrate to the US retail sector. We then ask: Quantitatively, how costly are policies that distort the size of production units? What is the impact of these policies on productivity? How do these policies affect the size distribution of establishments?

We find that policies that restrict the size of establishments in one of the sectors can generate nontrivial effects. In our calculations, productivity in the distorted sector can decline up to 25 percent, and average establishment size in the distorted sector can decline up to 50 percent. Furthermore, the policies we consider generate sizeable increases in the number of establishments in the distorted sector (up to 60 percent). This finding is a simple and strong implication of our framework, and is qualitatively consistent with observations from the Japanese case. Finally, even when the distorted sector in our calculations is relatively small (about 11 percent of total output in the absence of distortions), we find that the welfare cost of restrictions on size can be up to 0.7 percent. This leads us to conclude that policies of this sort are potentially costly.

The paper is organized as follows. Section 2 describes in some detail the regulations of the retail sector in Japan. Section 3 introduces the model economy we investigate. Section 4 discusses our choice of parameter values. Section 5 presents the findings from our experiments. Finally, Section 6 concludes.

2. Regulations on size in Japan: the case of the retail sector

Japan offers a unique and rather old case of protection of small retail shops. Owners of these shops constitute a strong pressure group, and as a result there exists national legislation that has aimed directly in the past, and indirectly in its present form, to protect and benefit them.

The origins of the regulations of large retail stores goes back to 1937, with the first ''Department Store Law'' enacted in reaction to complaints from small shop owners due to the expansion of large department stores. This law was eliminated in 1947 under the American administration, but was brought back under the same name in 1956. This law

stipulated a special procedure in order to get a license for the expansion of existing retail businesses, or the opening of new ones, beyond 1500 m^2.

The 1956 law applied to department stores, and thus other retail formats such as supermarkets, discount stores, etc., were not covered. As a result, the subsequent growth of these stores constituted a source of complaints for the retail lobby. Furthermore, the law focused on retail *businesses* of the department store category. This opened up a loophole under which large department stores were divided into separate business entities within the same building, each of them not exceeding 1500 m^2 (Larke, 1994). The complaints that this generated led to a major revision of the law, which took place in 1974. The new legislation, called Large Scale Retail Store Law, now focused on retail *stores*, closing thereby the loophole just described, and its scope was extended to include retail formats other than traditional department stores. The legislation specified an application process to get a license for retail stores above 3000 m^2 in big cities, and 1500 m^2 everywhere else.[2] In 1979, the law was reformed. The reform expanded severely the scope of the regulations under pressure of the retail lobby. It created two types of stores subject to restrictions, a model that continued until recently. Type-1 stores were those larger than 1500 m^2 (3000 m^2 in large cities), while Type-2 stores covered a group of a substantially small size: between 500 and 1500 m^2. Applications for stores of Type-1 were made to the Ministry of Trade and Industry (MITI), while applications for Type-2 were dealt at the local (prefectural) level.

The implementation of the law was altered in 1982, as the MITI introduced changes pertaining to stores of the first type. First, it provided local governments authority to restrict the opening of new stores in certain regions. Second, it created a new stage in the application process. This stage called for a concensus of interested parties, including those potentially affected by the opening (small, traditional stores). Notably, without concensus the whole process could not begin. The natural strategy of affected parties was not to provide concensus, as Larke (1994), p. 112, explains. As a result, most of the successful proposals for new stores in the 1980s took several years to complete.

By the mid-1980s, as a result of the law and the norms issued by the MITI governing its implementation, the process of obtaining approval for a new store at the Type-1 level was a long and costly one. It required a minimum of 7 different stages, and a maximum of 16. The first stage was a critical one, the local concensus stage, which could force the abandonment of the plans altogether. At many of these stages, the plans for the proposed new store could be stopped, or business plans could be forced to change by those negatively affected. It is worth noting that, most likely due to the increased severity and complexity of the regulations, the number of applications of the first type fell from about 399 in 1974 to about 157 in 1986; for Type-2 stores, the number of applications fell from 1029 in 1979 to about 369 in 1986.[3] To put these figures in perspective, it is worth emphasizing that the size of the

---

[2] An application had to specify at a minimum the proposed floor space, opening date, hours of operation, and the number of days in which the store would be closed during a year (see Ito, 1992 for details). By the early 1990s, the implementation of the law also set specified upper limits regarding closing times (07:00 p.m.), and a minimum number of annual closed days (44).

[3] *Source:* Larke (1994).

Japanese population is of about 120 million, and that the Japanese economy grew at an annualized rate of about 3.6 percent from 1974 to 1985.[4]

In 1992, the law was significantly relaxed for the first time. The most important change was the simplification of the application process, with the elimination of the first (concensus) stage, and a maximum of an year for the whole application process. Still, nonetheless, the lobby of small retailers retained a critical influence in the application process. Other changes included the increase in the lower limit for Type-1 stores to 3000 m^2 (6000 m^2 in big cities).

In 2000, the Large Scale Retail Location Law replaced the previous one. The new law requires the approval for stores larger than 1000 m^2, while the parties affected by the opening of a new store are still a critical part of the application process. The new legislation differs from the old one in two dimensions. First, all decisions are taken at the local level. Second, the protection of small retail is no longer an explicit objective of the legislation. The decision criteria now takes into account environmental factors (noise, congestion, etc.). It can be argued that the new legislation is even more restrictive than before. First, the limit on size now kicks in at 1000 m^2. This determines that while retail stores between 500 and 1000 m^2 are now unrestricted, all proposals for retail space above 1000 m^2 are considered locally. Second, as McKinsey-Global-Institute (2000) discusses, local governments are unlikely to see net benefits from a more competitive retail environment; these receive only a small share of their revenues from taxation of businesses as their operations are mostly financed from transfers from the federal government.

3. A two-sector model

We now describe a simple model with two production sectors and an endogenously determined size distribution of plants or establishments in each sector. The model is an extension of the Lucas (1978) span-of-control framework to multiple sectors. We first present the model economy without any distortion on size. We then proceed to put restrictions on size in one of the sectors.

The economy is inhabited by a single representative household. The household is comprised of a continuum of members of unit measure, who value two consumption goods, 1 and 2. The household values only consumption of these goods and maximizes

$$U(C_1, C_2), \tag{1}$$

where C_1 and C_2 denote the total household consumption of each good, respectively. The function $U(\cdot, \cdot)$ is continuous, strictly increasing in both arguments, strictly concave and differentiable.

Endowments. A fraction α of household members is of Type-1 and a fraction $1 - \alpha$ is of Type-2. A household member of type $i = 1, 2$ is endowed with z_i units of managerial ability. These efficiency units are distributed with support in $[0, \bar{z}]$ with cdf $F_i(z_i)$ and density $f_i(z_i)$.

---

[4] McCraw and O'Brien (1986) make a similar point.

Being of Type-1 implies that the household member can be a worker in any sector, or a manager in sector 1, with incomes that we describe below. Similarly, a household member of Type-2 can be a worker in any sector, or a manager in sector 2.

Production. Production of each good takes place, respectively, in sectors 1 and 2. Sector 1 produces good 1 while sector 2 produces good 2. We use, from now on, good 1 as the numeraire.

A manager in sector 1 has access to the technology

$$y_1 = A_1 z_1^{1-\gamma_1+\psi} n_1^{\gamma_1},$$

where $0 < \gamma_1 < 1$ and $\psi \geq 0$. Thus, production requires a managerial input (z_1) and labor (n). The manager maximizes profits taking the price of labor services as given, and obtains $\pi_1(z, w)$, which is the solution to

$$\max_n [A_1 z_1^{1-\gamma_1+\psi} n^{\gamma_1} - wn],$$

where w is the rental price for labor services. In a similar fashion, a manager in sector 2 has access to

$$y_2 = A_2 z_2^{1-\gamma_2+\psi} n^{\gamma_2}.$$

The manager maximizes profits and obtains $\pi_2(z, w, p)$, which is the solution to

$$\max_n [p A_2 z_2^{1-\gamma_2+\psi} n^{\gamma_2} - wn],$$

where p is the relative price of good 2 in terms of good 1.

The household problem. The problem of the household is to choose amounts of consumption goods 1 and 2, and the fractions of household members of each type who work as managers or workers in order to maximize utility.

If a household member becomes a worker, his/her efficiency units are transformed into 1 unit of labor and his/her income is then given by w. If instead he/she becomes a manager, his/her contribution to household's income is given by $\pi_1(z_1, w)$ or $\pi_2(z_2, w, p)$. Note that if both sectors are active, there exist unique thresholds $\hat{z}_1$ and $\hat{z}_2$ such that those individuals with efficiency units below the thresholds become workers, and those with efficiency units above the thresholds become managers. This follows from the fact that the functions $\pi_1(\cdot, w)$ and $\pi_2(\cdot, w, p)$ are strictly increasing and convex functions of the first argument (z) given diminishing returns to labor use.

Formally, the household problem is to select $\{C_1, C_2, \hat{z}_1, \hat{z}_2\}$ to maximize (1) subject to

$$C_1 + pC_2 = I(\hat{z}_1, \hat{z}_2, w, p),$$

where the income from managerial and labor services, $I(\hat{z}_1, \hat{z}_2, w, p)$, is given by

$$w[\alpha F_1(\hat{z}_1) + (1-\alpha)F_2(\hat{z}_2)] + \alpha \int_{\hat{z}_1}^{\bar{z}} \pi_1(z, w) f_1(z) \, dz$$

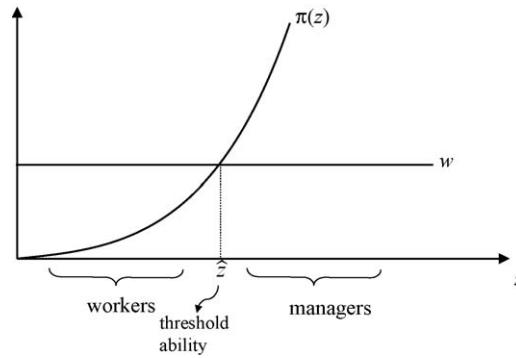$$+ (1-\alpha) \int_{\hat{z}_2}^{\bar{z}} \pi_2(z, w, p) f_2(z) \, dz.$$

Fig. 1. Occupational choice.

The solution to the household problem is characterized the following first-order conditions:[5]

$$\frac{U_2(C_1, C_2)}{U_1(C_1, C_2)} = p, \tag{2}$$

$$w = \pi_1(\hat{z}_1, w), \tag{3}$$

and

$$w = \pi_2(\hat{z}_2, w, p). \tag{4}$$

Condition (2) simply states that the marginal rate of substitution between both consumption goods must equal its relative price. Condition (3) states that the household member of Type-1 with marginal ability $\hat{z}_1$ must receive the same compensation as a manager than as a worker (i.e. be indifferent). The last condition (4) is the equivalent one for a household member of Type-2 with managerial ability $\hat{z}_2$. These indifference conditions defining occupational choice of household members are represented in Fig. 1.

Equilibrium. In equilibrium, the markets for labor services as well as the markets for both goods must clear. Let $n_1(z_1, w)$ be the demand for labor services of a manager of ability z_1 in sector 1. Similarly, let $n_2(z_2, w, p)$ be the demand for labor services of a manager of ability z_2 in sector 2. Market clearing in market for labor services requires

$$N^* = \alpha \int_{\hat{z}_1^*}^{\bar{z}} n_1(z, w^*) f_1(z) \, dz + (1 - \alpha) \int_{\hat{z}_2^*}^{\bar{z}} n_2(z, w^*, p^*) f_2(z) \, dz, \tag{5}$$

where an (*) over a variable denotes its equilibrium value, and N^*, aggregate labor supply, is given by:

$$N^* \equiv \alpha F_1(\hat{z}_1^*) + (1 - \alpha) F_2(\hat{z}_2^*).$$

Let $y_1(z_1, w)$ and $y_2(z_2, w, p)$ denote the supply of goods of 1 and 2 by managers with abilities z_1 and z_2, respectively. Then, market clearing for goods 1 and 2 requires:

$$\alpha \int_{\hat{z}_1^*}^{\bar{z}} y_1(z, w^*) f_1(z) \, dz = C_1^*. \tag{6}$$

---

[5] Given the structure of the problem, first-order conditions are necessary and sufficient.

and

$$(1 - \alpha) \int_{\hat{z}_2^*}^{\bar{z}} y_2(z, w^*, p^*) f_2(z)\, \mathrm{d}z = C_2^*. \tag{7}$$

We now summarize the presentation of the model with a definition of competitive equilibrium. A competitive equilibrium are values $\{C_1^*, C_2^*, \hat{z}_1^*, \hat{z}_2^*, w^*, p^*\}$, such that (i) given $\{w^*, p^*\}$, the values $\{C_1^*, C_2^*, \hat{z}_1^*, \hat{z}_2^*\}$ solve the household problem; (ii) the market for labor services clears (Eq. (5) holds); (iii) the markets for goods 1 and 2 clear (Eqs. (6) and (7) hold).

Discussion. A couple of implications of the framework are important to note at this point. First, since individuals of both types face the same wage rate as workers, the size of the smallest and average establishment can differ significantly across sectors. They depend critically on the parameters governing span-of-control and returns to managerial ability; γ_i and ψ.

Second, even if the smallest establishments in each sector differ, both sectors can have in equilibrium a positive mass of relatively large establishments. This model feature is key for our application of the model to the questions at hand. In the data, large establishments coexist with small ones in all sectors. Restrictions affecting size will tend to affect most severely potentially large establishments (that is, those run by the most able managers). Thus, to account for large establishments is important to reproduce features of the data and to assess the potential effects of policies on size.

3.1. Restrictions on size

Our representation of restrictions on size is meant to capture government policies which aim to affect the size of establishments via implicit taxes on input use. The central idea is that if an establishment wants to expand the use of an input beyond a given level, it faces a marginal cost of using the input in question that is larger than its market price.

We focus on restrictions imposed to large establishments in sector 2. We posit that the total cost associated to labor use beyond a pre-determined level $\underline{n}$ is given by

$$w\underline{n} + w(1 + \tau)(n - \underline{n}),$$

for some $\tau \in (0, 1)$. If $n \leq \underline{n}$, then the total cost of labor use is just $w\underline{n}$. Note that this resembles a progressive tax, in which there are two implicit marginal tax rates, 0 and τ. If $n > \underline{n}$, the production unit pays $w\underline{n}$ for the first $\underline{n}$ units used, plus an amount that is proportional to the difference between n and $\underline{n}$.

Our modeling of restrictions implies that the total cost associated to labor use is continuous in n. As a result, the function $\pi_2(\cdot)$ summarizing managerial rents, and the establishment demand function for labor services are continuous. Profit maximization dictates that there are potentially three types of establishments in the distorted sector. Unconstrained ones are small establishments that choose $n(z, w, p; \underline{n}, \tau) \leq \underline{n}$. Thus, for these establishments the marginal product of labor equals the rental rate w. On the other extreme, are those whose managers have relatively high levels of z, and thus choose $n(z, w, p; \underline{n}, \tau) > \underline{n}$. For these units, the marginal product of labor is *higher* than the wage rate. Finally, there is an intermediate group of establishments for which the marginal product of labor lies between w and $w(1 + \tau)$; for these, $n(z, w, p; \underline{n}, \tau) = \underline{n}$.
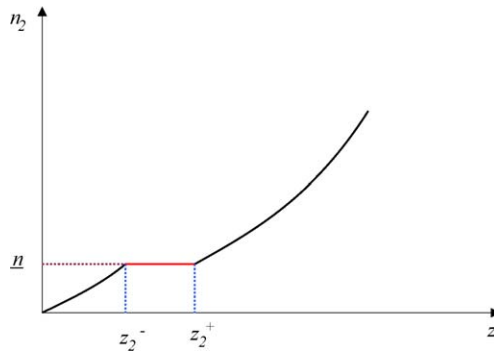
Fig. 2. The effects of restrictions on size.

Since the demand for labor services is continuous and increasing in managerial ability, the three types of establishments are mapped into levels of managerial ability. Hence, there exist thresholds z_2^- and z_2^+ so that: (i) unconstrained establishments are those with $z \in [\hat{z_2}, z_2^-)$; (ii) establishments in the intermediate group are those for which $z \in [z_2^-, z_2^+]$; (iii) the largest establishments have $z > z_2^+$. The resulting demand for labor services (as a function of managerial ability) is illustrated in Fig. 2.

We now briefly describe the modified household problem under restrictions on size. Resources taxed via restrictions on size are returned to the representative household in a lump-sum form. Formally, the household's budget constraint now equals:

$$C_1 + pC_2 = I(\hat{z}_1, \hat{z}_2, w, p; \underline{n}, \tau) + X,$$

where X stands for lump-sum transfers which are taken as given by the household. In equilibrium, these transfers amount to:

$$X^* = (1 - \alpha)\tau w^* \int_{z_2^{+*}}^{\bar{z}} (n(z, w^*, p^*; \underline{n}, \tau) - \underline{n}) f_2(z) \, dz.$$

4. Parameter values

We now choose parameter values in order to compute solutions to our model, which we do by selecting most of them so as to match a number of critical observations. To this end, we use data pertaining to the United States, which we take as a relatively distortion-free economy for the purposes of this paper.

We start by defining sector 2 as the retail sector defined in the National Income and Product Accounts (NIPA); sector 1 constitutes the rest of the economy, excluding the government sector. Based on these choices, we select parameter values as follows.

Preferences. We assume that the utility function takes the form:

$$U(C_1, C_2) = \log [H(C_1, C_2)],$$

where H is a C.E.S. aggregator, defined as:

$$H \equiv [\theta C_1^\rho + (1 - \theta) C_2^\rho]^{1/\rho}, \quad \rho \in (-\infty, 1).$$

In our benchmark case, we report results for the $\rho = 0$ (unitary elasticity of substitution), and later explore the implications of $\rho = -1/3$ (elasticity of substitution equal to 3/4) and $\rho = -2/3$ (elasticity of substitution equal to 3/5). We treat the parameter θ as an unknown, and choose its value so as to match the observed ratio of value added in the retail sector as a fraction of aggregate output, net of the government sector. This magnitude averaged about 11.0 percent for the period 1990–2000.[6]

Technology. We first set $A_1 = A_2 = 1$, and then proceed to select values for the degree of returns to scale in both sectors, γ_1 and γ_2, as well as the parameter ψ defining returns to managerial ability. To pindown these unknown parameters, we add three observations that the model is forced to match: the mean establishment size in the non-retail sector, the mean establishment size in the retail sector, and the fraction of workers in the labor force.

For the first two targets, we use the 1997 US Economic Census and calculate that the mean establishment size in the non-retail sector is of about 17.8 employees, while the corresponding mean value in the retail sector is of about 14.0 employees. Regarding the fraction of workers in the labor force, we target a value of 95 percent. We note that to pindown who is a worker and who is manager in actual data is difficult, and so we take this value as a compromise. From census data, it is possible to calculate a lower bound on the fraction of workers, as about 85.7 percent of the labor force performed non-managerial tasks in 2001.[7] Chang (2000), using PSID data, calculates an even lower value for the fraction of workers (84 percent). On the other extreme, a more literal interpretation of the model economy, which we prefer, suggests that each establishment is run by one manager. This consideration dictates a lower bound on the fraction of managers, which is obtained by dividing the number of active establishments in 1997 by the size of the work force in that year. This calculation leads to a fraction of workers in the population of about 96 percent.

Endowments. We assume that the distributions of potential managerial ability are log-normal and equal across sectors, so that $\log(z_i) \sim N(0, \sigma)$, $i = 1, 2$. In order to pindown σ and α, the fraction of individuals who have potential managerial abilities in sector 1, we add two observations relevant to the questions at hand. These are the dispersion in establishment size (in terms of workers), as measured by the coefficient of variation for both sectors. In the data, the distribution of establishment size is highly dispersed in both sectors, while both sectors display similar dispersion statistics. From the 1997 US Economic Census, we calculate that in the non-retail sector the coefficient of variation equaled 1.63, while in the retail sector the value for this statistic was 1.57.[8]

Summary. There are in total six parameters that we choose in order to reproduce observations. These are $\theta, \gamma_1, \gamma_2, \sigma, \alpha$ and ψ. Table 1 summarizes our choices. Table 2 lists the set of observations that constitute our targets, and shows the performance of the model in terms of them. The model has no problem in reproducing these targets, as the table demonstrates.

---

[6] *Source:* Economic Report of the President (2002), Table B-12. We use a 20 percent government to output ratio to calculate the ratio of value added to output we report.

[7] *Source:* Statistical Abstract of the US (2002), Table 588. This results from considering individuals under the occupation category "Executive, Administrative and Managerial".

[8] The high levels of dispersion in establishment size are hard not to emphasize. To put them in perspective, we note that this distribution is far more disperse than the distribution of labor earnings in the US, which has a coefficient of variation of about 0.7 (see Haider, 2001 for instance).

Table 1
Parameter values

Parameter	Value
γ_1	0.746
γ_2	0.709
σ	2.135
θ	0.89
ψ	0.0875
α	0.875

Table 2
Targets

Statistic	Data	Model
Mean size		
Sector 1	17.8	17.9
Sector 2	14.0	14.0
Coeff. variation		
Sector 1	1.63	1.62
Sector 2	1.57	1.58
Fraction workers	0.95	0.95
Value added sector 2 (% GDP)	0.11	0.11

5. Findings

We now conduct experiments to quantitatively evaluate the impact of restrictions on size. We proceed by comparing equilibria across different size restrictions. We report results for restrictions at three levels. In the first case, $\underline{n}$ equals average labor use in sector 2 without restrictions. In the second and third case, distortions are more severe, and $\underline{n}$ is equal to two-thirds of average labor use and to median labor use in sector 2, respectively, in the absence of restrictions. In both cases, we report results for a relatively low value of the implicit tax rate ($\tau = .20$), and for a relatively high value ($\tau = .50$).

Aggregates. Table 3 summarizes the main findings for aggregate variables. Output in the distorted sector drops from 6.3 percent to 16 percent; the magnitude in the fall depends on the interplay between the location of the distortion in the size distribution, and the increase in the magnitude of the implicit tax rate, τ. As τ increases, affected establishments either set their demand for labor services at $\underline{n}$, or demand labor services from a new, higher price $w(1 + \tau)$. This process leads to a reduction in the total demand for labor services of the sector, and to a reduction in the overall supply of the good produced by the distorted sector. In equilibrium, the undistorted sector expands, the relative price of the good produced in the distorted sector increases, which is accompanied by an increase in the number of small establishments in this sector as Table 3 shows. It is worth noting the phenomenon that total output of the distorted sector decreases, despite the emergence of new, small establishments; this reflects the fact that large (distorted) ones account for a disproportionate share of the total supply of the good in question.

Table 3
Aggregate and productivity effects

Statistic	Implicit tax rate		
	0	0.2	0.5
n = Mean labor			
Aggr. output	100	100.6	100.9
Aggr. output[a]	100	99.9	99.5
Output sector 2	100	93.7	88.5
Output per-worker sector 2 (a)	100	96.4	93.3
Output per-worker sector 2 (b)	100	89.2	81.0
Number establishments sector 2	100	119.7	138.0
n = (2/3) Mean labor			
Aggr. output	100	100.8	101.2
Aggr. output[a]	100	99.9	99.5
Output sector 2	100	92.8	86.2
Output per-worker sector 2 (a)	100	96.9	93.6
Output per-worker sector 2 (b)	100	87.8	77.7
Number establishments sector 2	100	122.5	147.9
n = Median labor			
Aggr. output	100	100.9	101.4
Aggr. output[a]	100	99.9	99.5
Output sector 2	100	91.9	84.0
Output per-worker sector 2 (a)	100	98.1	95.1
Output per-worker sector 2 (b)	100	86.2	74.5
Number establishments sector 2	100	126.8	159.2

[a] At benchmark (undistorted) prices.

We emphasize that the increase in the number of small establishments is a simple and natural implication of our framework, which is qualitatively consistent with the observations pertaining to the Japanese retail sector we discussed earlier. Quantitatively, the increase in the number of small establishments is rather substantial, ranging from about 20 percent to 59 percent.

Productivity. The restrictions on size have a direct and negative impact on productivity measures. We report, in Table 3, two different ones. The first notion of productivity is simply average output per-unit of labor across establishments in sector 2: we label it output per-worker (a). The second one, aggregate output per-unit of aggregate labor utilized in sector 2, when labor includes efficiency units supplied by managers: we label it output per-worker (b). Formally,

$$\frac{Y_2}{N_2 + Z_2},$$

where Y_2, N_2 and Z_2 stand for total output, labor and managerial efficiency units in sector 2, respectively.

Both notions of productivity drop as restrictions are introduced. The drop in the first notion of productivity ranges from 1.9 percent to 6.7 percent. When efficiency units of managers are taken into account in the notion of labor, the fall is much sharper: from 10.8

percent to 25.5 percent. This reflects the low output of the small establishments that emerge as restrictions are introduced.

It is critical to understand why the drop in productivity occurs. We focus now in detail on the case of output per-worker, as this is a statistic usually computed in empirical studies. Physical output per-worker (y_2/n_2) in a non distorted establishment equals:[9]

$$\frac{w}{p\gamma_2}.$$

Therefore, whenever the ratio of wage rates to the relative price falls, this notion of productivity drops in undistorted establishments. For establishments demanding labor services above $\underline{n}$, output per-worker amounts to:

$$\frac{w(1+\tau)}{p\gamma_2}.$$

Thus, output per-worker in establishments in the latter group increases relative to the undistorted case whenever the effect of the implicit tax τ dominates the effect of changes in the ratio of w to p. This turns out to be case in our economy. Nevertheless, as the fraction of distorted establishments falls (strongly) as restrictions are introduced, while the opposite occurs for undistorted ones, average output per-worker falls when restrictions are introduced as Table 3 shows.

We note that the simple reasoning above has important implications for measurement. If output at distorted prices is used to calculate output per-worker (i.e. py_2/n_2), the drop in this statistic might not be detected. In the context of our simple economy, relative prices would cancel out and all variation in output per-worker at distorted prices would be associated to movements in wage rates. If undistorted prices are used instead, the drop in output per-worker measured in physical units that we report is equivalent to a drop in output per-worker, when output is measured at undistorted prices.

To what extent output per-worker would fall if restrictions are placed on inputs other than labor? Suppose production requires only two inputs, capital and labor, and technology assumes the form used in many studies, $y_i = A_i z^{1-\gamma_i+\psi} g(k,n)^{\gamma_i}$, $i = 1, 2$, where $g(\cdot, \cdot)$ displays constant returns to scale. Assume further that size restrictions hinge upon the use of capital services above a threshold, say $\underline{k}$. In this case, it is easy to show that in the distorted sector the capital to labor ratio is a continuous, weakly decreasing function of z, taking a relatively high value for establishments demanding $k < \underline{k}$, and a relatively low one for those demanding $k > \underline{k}$, while strictly decreasing with z for those demanding $\underline{k}$. Furthermore, if the function $g(\cdot, \cdot)$ is Cobb–Douglas, then output per-worker in sector 2 is constant across all establishments, distorted and undistorted, and equal to:

$$\frac{w}{p(1-v)\gamma_2}.$$

where $(1 - v)$ is the labor share in $g(\cdot, \cdot)$. Therefore, as the implicit tax does not appear in the expression above, a given change in the ratio of w to p generated by size restrictions

---

[9] To calculate output per-worker, we first calculate the optimal demands for inputs for each establishment, and then obtain the respective supply function, $y_2(z, w, p)$. We then obtain physical output per-worker as $y_2(z, w, p)/n(z, w, p)$.

Table 4
Effects on size distribution

Statistic	Implicit tax rate		
	0	0.2	0.5
$\underline{n}$ = Mean labor			
Mean size sector 2	14.01	10.67	8.78
Coeff. variation sector 2	1.58	1.28	0.86
Median size sector 2	5.93	6.10	6.30
% Distorted ($n \geq \underline{n}$)	23.2	24.5	25.6
% Distorted ($n < \underline{n}$)	23.2	13.4	6.4
$\underline{n}$ = (2/3) Mean labor			
Mean size sector 2	14.01	10.21	7.92
Coeff. variation sector 2	1.58	1.38	0.96
Median size sector 2	5.93	6.15	6.40
% Distorted ($n \geq \underline{n}$)	33.8	35.3	37.0
% Distorted ($n < \underline{n}$)	33.8	20.1	10.3
$\underline{n}$ = Median labor			
Mean size sector 2	14.01	9.75	7.03
Coeff. variation sector 2	1.58	1.50	1.17
Median size sector 2	5.93	5.93	5.93
% Distorted ($n \geq \underline{n}$)	50.0	51.6	53.3
% Distorted ($n < \underline{n}$)	50.0	30.6	16.6

leads to a larger reduction in output per-worker than in the case where labor is the only input. From these perspectives, the productivity figures associated to size restrictions that we report in Table 3 are conservative estimates.

Size distribution effects. Table 4 shows key statistics related to the effects of restrictions on the size distribution of establishments, and shows that they have rather substantial consequences on it. The mean size of establishments is very sensitive to size restrictions: it ranges from about 10.7 to 7.0 with restrictions in place, while it is about 14.0 in the absence of them. That is, the value of this statistic falls between 23.6 percent and 50 percent. In contrast, the size of the median establishment moves in the opposite direction, except when restrictions kick-in at median labor use. This occurs in spite of the appearance of small establishments at the bottom of the distribution. The expansion of undistorted establishments as restrictions are introduced accounts for this.

Dispersion in size, measured by the coefficient of variation, drops as Table 4 indicates. It is worth mentioning that several forces influence the dispersion in the size distribution. On the one hand, everything else constant, the emergence of new, small establishments tend to increase dispersion. On the other hand, the reduction in the size of distorted establishments contribute to reduce dispersion, while the increase in size of undistorted ones has an uncertain effect. Overall, the effects that lead to a reduction in dispersion dominate, as the results show.

It is important to notice the effects that restrictions have upon the mass of establishments at or above $\underline{n}$, the level where these restrictions kick-in. In the first place, note that the restrictions create a sizeable mass of establishments concentrated at $\underline{n}$; the mass of establishments at this level jumps from theoretical level of zero in the undistorted case, to values ranging from 11.1 percent to 36.7 percent. Both the

Table 5
Welfare costs

Distortion location	Implicit tax rate	
	0.2	0.5
n = Mean labor	0.14	0.55
n = (2/3) Mean labor	0.14	0.61
n = Median labor	0.15	0.67

contraction of some distorted establishments, which now demand labor services at $\underline{n}$, and the expansion of previously undistorted ones account for this phenomenon. Second, the relatively severe increase in the implicit tax rate from 20 percent to 50 percent does *not* change significantly the overall mass of establishments affected by the restrictions. It is worth emphasizing that this phenomenon can lead to an erroneous conclusion, such as that an increase in the severity of the restrictions does not matter. To see this, notice that the increase in the implicit tax rate leads to a significant decrease in the number of establishments strictly above $\underline{n}$. Quantitatively, when τ increases, this magnitude drops from 23.2 percent to 13.4 percent when $\underline{n}$ is equal to average labor use in sector 2 without distortions, and from 50.0 percent to 36.7 percent when $\underline{n}$ is equal to median labor use in the sector.[10]

Welfare. We now look at the welfare costs associated with the policies we investigate. We calculate the associated welfare cost as the percentage increase in consumption of both goods that is necessary in order to make the representative household indifferent across distorted and undistorted equilibria. That is, we find the value of Δ that solves:

$$V^* = U(C_{1d}(1 + \Delta), C_{2d}(1 + \Delta)),$$

where V^* is the realized utility level in the undistorted equilibrium, and C_{1d} and C_{2d} are the consumption of good 1 and good 2 in equilibria with distortions. Table 5 shows that welfare costs associated with these policies can be significant; they range in our exercises from a relatively low value of 0.14 percent to a much higher one of 0.67 percent. The relatively large decline in the consumption of good 2 reported previously, which is only partially mitigated by the increase in consumption of good 1, is responsible for these welfare costs. These figures are in the ballpark of existing welfare cost estimates derived from static models—see Shoven and Whalley (1984) for a comprehensive review.

It is important to place the welfare results in perspective. First, how large are the distortions we impose on our model economy? Surprisingly, they are relatively small. Note from Table 2 that sector 2 accounts for only 11 percent of total output. Moreover, in our experiments only about 23.2 percent to 50 percent of the establishments in sector 2 are affected by size restrictions, and only about 6.4 percent to 30.6 percent of the establishments effectively pay the implicit tax on labor services. Finally, the establishments that pay this tax, only pay a penalty on the amount of labor services they rent above the threshold level, $\underline{n}$. Indeed, one can calculate in this economy the total value of tax payments

---

[10] Note that when restrictions kick-in at the median value of labor use in undistorted case, median establishment size is does not change with restrictions.

as a percentage of total payments for labor services in sector 2. This calculation gives an average tax rate on payments to labor in sector 2, which equals:

$$
\frac{\tau \int_{z_2^{+*}}^{\bar{z}} (n_2(z, w^*, p^*; \underline{n}, \tau) - \underline{n}) f_2(z) \, dz}{\int_{z_2^*}^{\bar{z}} n_2(z, w^*, p^*; \underline{n}, \tau) f_2(z) \, dz}.
$$

In our experiments, this average tax rate turns out to be relatively small. It ranges from 5.6 percent when $\tau = 0.2$ and $\underline{n}$ is equal to mean level of labor used in sector 2, to 14.9 percent when $\tau = 0.5$ and $\underline{n}$ is equal to median labor use in sector 2. That is, the highest average tax rate turns out to be smaller than the lowest implicit tax rate we use (20 percent). Although this average tax rate in sector 2 is low, the underlying implicit tax affects the decisions at the margin of large establishments. These establishments account for the bulk of output in the sector: in the undistorted economy, establishments above the median size are responsible for about 86.5 percent of total output of sector 2, while establishments above the mean account for 69.4 percent.

Second, in order to further gauge the importance of the implicit taxes on the decisions of large establishments, we compute welfare costs when we impose the highest average tax rate of 14.9 percent on *all* establishments in sector 2. The resulting welfare cost equals 0.067 percent, lower than any of the welfare costs reported in Table 5!

Finally, for the case of Japan when we are guided by observations from its regulated retail sector, we conclude that the cost of the policies in place is closer to the upper end of estimates reported in Table 5 than otherwise. We note that the mean size of Japanese retail establishments is of about 6.4 employees.[11] This is less than half the mean size of American retail establishments (about 14). While there are factors other than size restrictions that affect the size of retail establishments,[12] we emphasize that the observed mean size is below any of the corresponding ones implied by the policies we study.

The role of the elasticity of substitution. In our benchmark economy, we assumed a unitary elasticity of substitution between two consumption goods and set $\rho = 0$. In this section, we revisit the effects of restrictions on size when two goods are less substitutable. In the absence of empirical estimates of this elasticity, we report results below for a value of it less than one (0.75); there are a number of reasons to suspect that the degree of substitution in preferences between retail and non-retail consumption goods is low. We follow the same procedure we used before to select the parameters: we choose the six parameters ($\gamma_1, \gamma_2, \sigma, \psi, \theta$, and α), so as to match the same six targets in Table 2.[13]

In Table 6, we report the effects of restrictions on size when $\underline{n}$ equals the average labor use in sector 2 without any size restrictions. The basic picture that emerges from this Table (compared to Tables 3 and 4) is that a lower elasticity of substitution, when other parameters are adjusted to reproduce targets, only moderately changes the consequences on the variables of interest. Output in sector 2 declines by 5 percent (instead of 6.3 percent) when $\tau = 0.2$, and by 9.5 percent (instead of 11.5 percent) when $\tau = 0.5$. The number of establishments in sector 2 increases by 22.2 percent (instead of 19.5 percent) when $\tau = 0.2$,

[11] *Source:* Japan's 2001 Enterprise and Establishment Census. Available at http://www.stat.go.jp/english/data/jigyou/index.htm.

[12] See Flath (2003) for a discussion.

[13] The values are now $\gamma_1 = 0.743$, $\gamma_2 = 0.711$, $\sigma = 2.145$, $\theta = 0.938$, $\psi = 0.089$ and $\alpha = 0.865$.

Table 6
Effects with $\rho = -1/3$ ($\underline{n}$ = mean labor use)

Statistic	Implicit tax rate		
	0	0.2	0.5
Aggr. output	100	100.7	101.0
Aggr. output[a]	100	99.9	99.5
Output sector 2	100	95.0	90.5
Output per-worker sector 2 (a)	100	96.0	92.6
Output per-worker sector 2 (b)	100	90.6	83.2
TFP sector 2	100	91.4	84.6
Number establishments sector 2	100	122.2	143.1
Mean size sector 2	14.00	10.72	8.86
Coeff. variation sector 2	1.56	1.27	0.85
Median size sector 2	5.96	6.18	6.40
% Distorted ($n \geq \underline{n}$)	23.3	24.7	26.0
% Distorted ($n < \underline{n}$)	23.3	13.5	6.5

[a] At benchmark (undistorted) prices.

Table 7
Welfare costs with $\rho = -1/3$ (% increase in consumption)

Distortion location	Implicit tax rate	
	0.2	0.5
$\underline{n}$ = Mean labor	0.16	0.58
$\underline{n}$ = (2/3) Mean labor	0.16	0.65
$\underline{n}$ = Median labor	0.17	0.69

and by 43.1 percent (instead of 38 percent) when $\tau = 0.5$. Similarly, the decline in the average establishment size in sector 2 is slightly higher. From these results, we conclude that a reduction in the elasticity of substitution between the goods, when the rest of the parameters are adjusted to reproduce observations, does not change the quantitative findings we reported previously in a fundamental way.[14]

When goods are more difficult to substitute in consumption relative to the benchmark case, the distorted sector attracts more resources as size restrictions are introduced. As a result, size restrictions have in equilibrium a larger effect on the relative price of good 2. When $\underline{n}$ is at the average level of labor use in an economy without any restrictions on size and $\rho = 0$, the relative price of good 2 rises by 7.1 percent with $\tau = 0.2$, and by 14.2 percent with $\tau = 0.5$. The corresponding numbers when $\rho = -1/3$ are 7.9 percent and 15.3 percent, respectively. A larger price increase leads to a larger decline in output per-worker in sector 2. It also causes a larger increase in the number of small establishments, by making the production of the second sector more attractive for these. Not surprisingly, we also get slightly larger welfare effects as Table 7 demonstrates.

---

[14] We have also experimented using lower values for the elasticity of substitution. When $\rho = -2/3$, the predictions are very similar to the case of $\rho = -1/3$. Output in sector 2 falls by 4.3–8.9 percent, mean size declines to 10.8 and 8.9 employees, and the number of establishments in sector 2 increases by 21.9–43.8 percent.

6. Conclusion

In this paper, we analyze government policies that target establishments of different sizes. To this end, we develop a two-sector model economy in which agents differ in terms of their sector-specific skills, and sort themselves into managers and workers. We interpret these two sectors as the retail and the remaining sectors, and calibrate our benchmark economy to be consistent with observations from the US economy. We then consider policies that increase factor prices for larger establishments in the retail sector.

We find that these policies can have potentially large effects. Our simulations show that such policies reduce output per-worker in the distorted sector, while leading to a significant increase in the number small establishments. We view these results as consistent with observations on Japanese retail sector—a sector with strict size regulations. Our simulations also show that these policies can generate significant welfare losses. The presence of large establishments which accounts, both in the model and in the data, for a disproportionate large fraction of output in each sector, plays a key role in these results.

Acknowledgments

References

Baily, M.N., 1993. Competition, regulation, and efficiency in service industries. Brookings Papers on Economic Activity 71–159.

Baily, M.N., Solow, R.M., 2001. International productivity comparisons built from the firm level. Journal of Economic Perspectives 15, 151–172.

Basu, M., Fernald, J., Oulton, N., Srinivasan, S., 2003. The case of missing productivity growth: or, does information technology explain why productivity accelerated in the United States but not in the United Kingdom. NBER Working Paper 10010.

Bergoeing, R., Kehoe, P., Kehoe, T., Soto, R., 2002. A decade lost and found: Mexico and Chile in the 80's. Review of Economic Dynamics 5, 166–205.

Chang, Y., 2000. Wages, business cycles and comparative advantage. Journal of Monetary Economics 46, 143–171.

Council of Economic Advisers, 2002. Economic Report of the President. United States Government Printing Office, Washington, DC.

Flath, D., 2003. Regulation, distribution efficiency, and retail density. NBER Working Paper 9450.

Gollin, D., 1995. Do Taxes on Large Firms Impede Growth? Evidence from Ghana. Economic Development Center, Paper 95–9, University of Minnesota.

Haider, S., 2001. Earnings instability and earnings inequality of males in the United States: 1967–1991. Journal of Labor Economic 19 (4), 799–836.

Ito, T., 1992. The Japanese Economy. MIT Press, Cambridge.

Larke, R., 1994. Japanese Retailing. Routledge, London and New York.

Little, I., Mazumdar, D., Page, J.M., 1987. Small Manufacturing Enterprises: A Comparative Analysis of India and Other Economies. Oxford University Press for the World Bank.

Lucas, R.E., 1978. On the size distribution of business firms. Bell Journal 9, 508–523.

McCraw, T., O'Brien, P., 1986. Production and distribution: competition policy and industry structure. In: McCraw, T. (Ed.), America vs. Japan. Harvard Business School Press.

McKinsey-Global-Institute, 2000. Why the Japanese Economy is not Growing: Microbarriers to Productivity Growth. McKinsey Global Institute.

Nicoletti, G., Scarpetta, S., 2003. Regulation, productivity and growth: OECD evidence. OECD Economics Department Working Paper, 347.

Parente, S., Prescott, E., 2000. Barriers to Riches. MIT Press, Cambridge.

Restuccia, D., Rogerson, R., 2003. Policy Distortions and Aggregate Productivity with Heterogeneous Plants. University of Toronto, Mimeo.

Schmitz, J., 2001. Government production of investment goods and aggregate labor productivity. Journal of Monetary Economic 47, 163–187.

Shoven, J. Whalley, J., 1984. Applied general equilibrium models of taxation and international trade: an introduction and survey. Journal of Economic Literature 22, 1007–1051.

US Census Bureau, 2002. Statistical Abstract of the United States, 2002, 122nd ed. Washington DC.