SERIEs (2019) 10:401–418
https://doi.org/10.1007/s13209-019-00209-w

ORIGINAL ARTICLE

Tax efficiency in a model of endogenous markups

Esra Durceylan[1]

Received: 15 May 2019 / Accepted: 11 October 2019 / Published online: 30 October 2019

Abstract
Efficiency comparison of ad valorem and unit taxes has been traditionally based on consumer welfare. However, if the tax instrument also affects the distribution of firms over their productivities, the policy maker may be concerned about the implications on aggregate productivity as well. This paper makes an efficiency comparison of ad valorem and unit taxes by allowing the distribution of firms to respond to changes in policy. First, I make an efficiency comparison in a model with monopolistically competitive firms that are homogenous with respect to their productivity levels. Consumer preferences exhibit love for variety and allow firms to adjust their markups. I find that ad valorem tax is more efficient. Allowing for firm heterogeneity overturns this result at high revenue requirements. As the tax rate increases, ad valorem tax causes excessive exit of firms which makes the market more competitive. Hence, few surviving firms price lower by decreasing their markups. Lower prices decrease the tax revenue collected. As a result under ad valorem tax regime, higher consumer surplus is dominated by lower tax revenue. On the other hand, production is concentrated among relatively more productive firms. Thus, aggregate productivity is higher under ad valorem tax regime.

Keywords Unit tax · Ad valorem tax · Efficiency · Monopolistic competition · Heterogenous firms

JEL Classification L11 · H21 · H25

1 Introduction

Governments tax consumption mainly for two reasons. First, they want to raise revenue to finance spending on public goods. Second, they use taxation to limit negative externalities that arise from consumption or production of certain goods. Such taxes

✉ Esra Durceylan
esradk@sabanciuniv.edu

[1] Faculty of Arts and Social Sciences, Sabanci University, Orta Mh. Universite Cd. 27B, 34956 Istanbul, Turkey

are known as corrective taxes. Taxing tobacco products, alcoholic beverages, and fuel are some well-known examples. Corrective taxes are implemented as unit taxes or ad valorem taxes. The UK government, for example, taxes gambling revenue via ad valorem taxes, whereas it taxes alcohol and tobacco with unit taxes.[1] Similarly, in March 1995, the UK government published a paper on landfill tax and proposed ad valorem taxation rather than a unit tax charged by weight arguing the balance between environmental concern and practical feasibility. However, the government introduced the landfill tax as a unit tax per weight in August 1995 (Seely 2009). Which one is more efficient? The choice of these taxes can in many ways affect their impact on the industry. Both types of taxes distort the pricing behavior of different firms differently. The type of taxes chosen by the policy maker not only affects consumer welfare, it also affects the aggregate productivity in an industry by changing the composition of surviving firms.

In this paper, I study the efficiency ranking of output equivalent unit and ad valorem taxes in a model of monopolistically competitive firms that produce differentiated products. A key feature of the model is that firms adjust their markups in response to changes in the economic environment. De Loecker et al. (2019) calculate markups of all publicly traded firms covering all sectors of the US economy over the period 1955–2016. They find that not only markups are not constant within an industry, but also the distribution of markups has become more skewed over time. They conclude that policy makers have to be cautious not to use the average markup as that of a representative firm to draw conclusions about the economy as a whole.[2]

I compare the tax efficiencies in two possible settings. First, with homogenous firms, I show that ad valorem tax is more efficient than an output equivalent unit tax. It generates more tax revenue and consumer surplus. Firms are larger (hence less variety) but they price lower. Hence, a firm's markup is relatively lower under ad valorem tax regime. Allowing firms to change their markups plays a key role in this finding and overturns the superiority of unit tax under almost identical model setting with constant markup pricing by Dröge and Schröder (2009).

Next, I introduce firm heterogeneity to the model. Consumer surplus is higher under ad valorem tax regime. However, with heterogenous firms, unit tax becomes a better instrument to collect higher tax revenues. Since ad valorem tax causes the marginal surviving firm to be more productive, the market is relatively more competitive under this regime. As a result, highly productive firms survive with relatively larger output levels and lower prices. This comes at an expense of lower tax revenues. Firms with relatively higher (lower) levels of productivity choose higher (lower) levels of markups under unit taxes than under ad valorem taxes. Moreover, as the tax rates increase, these gaps expand. Hence, allowing for endogenous markups augments the difference in tax revenue and number of varieties under the two tax regimes in favor of unit tax when the tax requirements are high.

As the existing studies state, the choice of taxes poses a trade-off between revenue extraction and variety generation. On one hand, ad valorem tax is a better revenue

[1] Source: www.gov.uk/tax.

[2] There is a growing body of literature in international trade that recalculates gains from trade with variable markups (Arkolakis et al. 2019; Edmond et al. 2015).

extractor. On the other hand, a unit tax generates a greater variety due to the larger tax overshift and thus enables firms to survive at lower output levels. In addition to this trade-off, if one adds firm heterogeneity into these models, these taxes have nontrivial effects on the selection of firms and thus on aggregate productivity. Since unit taxation allows firms to survive at lower levels of output, the marginal surviving firm under unit tax regime is less efficient than the marginal surviving firm under ad valorem tax regime. Moreover, resource allocation is more efficient under ad valorem tax regime since relatively more productive firms have higher market shares. In addition to these forces, this paper contributes to this literature by arguing that allowing for firms to adjust their markups plays a key role in the efficiency ranking of these taxes. First, firms are able to pass the tax burden onto consumers better under unit taxes. Second, with endogenous markups, the tax regimes cause different firms to pass the burden to consumers differently.

Majority of the existing studies that compare efficiencies of these taxes use the equal yield criterion. Early contributions state that in a perfectly competitive market, the choice of tax policy does not matter; since firms are price takers, only the cost-price increase generated by the tax is relevant. Consequently, for every unit tax rate, there exists an equivalent ad valorem tax level. On the other extreme, when there is a monopoly, this equivalence result breaks down: an ad valorem tax generates more welfare than a unit tax [see Suits and Musgrave (1953) and Skeath and Trendel (1994)]. The result is driven by the fact that a profit-maximizing monopolist increases its output when an ad valorem tax is imposed. The superiority of ad valorem tax is shown to hold for a wide range of market structures.[3] In a monopolistic competition setting, Schröder (2004) with Dixit–Stiglitz preferences that result in constant markup pricing shows that an ad valorem tax remains welfare superior. This result is robust to add cost asymmetries to this setting (Schröder and Sørensen 2010). This paper also contributes to this literature by numerically showing that efficiency ranking of revenue equivalent ad valorem and unit taxes depend on the level of tax rates if one accounts for markup variation. In line with the intuition above, endogeneity of markups leads to reversal of the efficiency ranking as tax rates increase.

The literature on the relative efficiency of corrective taxes uses equal output criterion as most of the time, the size of the externality is correlated with the size of output or consumption. For perfectly competitive markets, the two regimes generate identical outcomes. In a Cournot setting, if the externality is small, ad valorem taxation is optimal. If the externality is large enough, then unit taxation becomes a better instrument (Pirttilä 2002). Dröge and Schröder (2009) show that unit tax regime is optimal if the market is composed of monopolistically competitive homogenous firms. This paper contributes to this literature by showing that, first, accounting for markup variation overturns the superiority of unit taxation shown by Dröge and Schröder (2009). Second, if one further incorporates firm heterogeneity along with endogenous markup, the parameter-independent results are challenged. At higher (lower) tax requirements, unit (ad valorem) tax regime is better at maximizing welfare.

The remaining of this paper is organized into four additional sections. Section 2 details consumer preferences, demand for differentiated good and labor market. Sec-

[3] See Keen (1998) for a review of the literature.

tion 3 explains supply side of the market populated with homogenous firms under these two tax regimes. Section 4 introduces heterogeneity of firms in cost of production. Finally, Sect. 5 concludes.

2 Model

The model used in this paper is based on Melitz and Ottaviano (2008). There are L identical consumers of size one. Their preferences are defined over a continuum of differentiated products and a homogenous good. They choose differentiated products from the set of varieties, Θ. The homogenous good is indexed by 0, and the differentiated good is indexed by i. A consumer's utility function is

$$U = q_0 + \alpha \int_{i\in\Theta} q_i \mathrm{d}i - \frac{1}{2}\gamma \int_{i\in\Theta} (q_i)^2 \mathrm{d}i - \frac{1}{2}\eta \left(\int_{i\in\Theta} q_i \mathrm{d}i \right)^2, \tag{1}$$

where q_0 is homogenous good consumption and q_i is differentiated good consumption of variety i. α, η and γ are the positive demand parameters. The degree of product differentiation increases with γ and $\int_{i\in\Theta}(q_i)^2 \mathrm{d}i$ reflects consumers' love for variety.

In this economy, the demand for the homogenous good is assumed to be positive. The technology to produce the homogenous good is identical across firms. It is produced under constant returns to scale at a cost of one unit of labor and its price is normalized to one.

Utility maximization gives a consumer's inverse demand for each differentiated variety as

$$p_i = \alpha - \gamma q_i - \eta Q, \tag{2}$$

where $Q = \int_{i\in\Theta} q_i \mathrm{d}i$ is the total quantity consumed. When I aggregate over all consumers, the linear market demand for variety i is given as

$$q_i = \frac{\alpha}{\gamma + \eta N} - \frac{1}{\gamma} p_i + \frac{\eta N}{\gamma + \eta N} \frac{1}{\gamma} \widetilde{P},$$

where N is the mass of differentiated varieties, and $\widetilde{P} = \frac{1}{N} \int_{i\in\Theta} p_i \mathrm{d}i$.

Consumers are endowed with one unit of labor. Labor is inelastically supplied, and there is no leisure in the utility function. Furthermore, labor is the only factor of production. Homogenous good is freely traded in a perfectly competitive market. Since workers receive their revenue of marginal product which is equal to the price of the homogenous good, the wage is unity.

The model features of consumers preferences are identical under both homogenous and heterogenous cost cases. Going forward, I lay out remaining features of the model that differ for differentiated good producers when they have either homogenous or heterogenous costs, and when they pay either ad valorem or unit taxes.

3 Relative efficiency of tax regimes with homogenous costs

Firms produce differentiated goods at a constant marginal cost of c.[4] These firms must pay a sunk entry cost of f_e in order to enter the market.[5] There is free entry and exit. Firms therefore enter the market if the expected profit of a firm covers the sunk entry cost. This condition determines the total mass of firms and variety in the market and is called the free entry condition.

Ad valorem tax regime

I first analyze the market if the government collects ad valorem taxes as fraction t of price of a good sold in the market. In this economy, a profit-maximizing firm chooses an optimal amount of output given the inverse demand it faces. A firm supplies $q_t(c)$ units given as

$$q_t(c) = \frac{1}{\gamma}\left(p_t(c) - \frac{c}{1-t}\right), \tag{3}$$

where going forward subscript t refers to ad valorem taxing economy variables. The free entry condition of firms is given as

$$\pi_t(c) - f_e = 0.$$

The equilibrium in the market is characterized by profit-maximizing price, output by each firm and the total number of varieties in the market given as

$$\begin{aligned} q_t(c) &= \sqrt{\frac{f_e}{\gamma(1-t)}}, \\ p_t(c) &= \sqrt{\frac{f_e\gamma}{(1-t)}} + \frac{c}{1-t}, \\ N_t &= \frac{Q_t}{q_t(c)} = \frac{\alpha - \frac{c}{1-t} - 2\sqrt{\frac{\gamma f_e}{(1-t)}}}{\eta\sqrt{\frac{f_e}{\gamma(1-t)}}}. \end{aligned} \tag{4}$$

Unit tax regime

In this environment, government collects s for every unit sold in the market. As a result, firms's profit-maximizing output is

$$q_s(c) = \frac{1}{\gamma}\left(p_s(c) - (c+s)\right), \tag{5}$$

where going forward subscript s refers to unit taxing economy variables. The free entry condition is characterized similarly to that of an ad valorem tax as $\pi_s(c) - f_e = 0$.

[4] There is no need to add a fixed cost of production in order to generate a finite mass of firms in the market. The linear demand caps the maximum price a firm can charge in order to make positive profit.

[5] Since wage is unity, all cost figures mentioned in this paper are in unit labor requirements.

The equilibrium values of profit-maximizing price, quantity and equilibrium number of varieties are

$$\begin{aligned} q_s(c) &= \sqrt{\frac{f_e}{\gamma}}, \\ p_s(c) &= \sqrt{f_e\gamma} + c + s, \\ N_s &= \frac{Q_s}{q_s(c)} = \frac{\alpha - c - s - 2\sqrt{\gamma f_e}}{\eta\sqrt{\frac{f_e}{\gamma}}}. \end{aligned} \tag{6}$$

Welfare is calculated as the sum of consumer surplus, tax revenue and producer surplus in this economy. The producer surplus is zero since there is free entry and exit. Since tax revenue is redistributed back to consumers, the calculated welfare per person is identical to the utility [expression (1)] derived from consuming all the varieties in the market.

Proposition 1 *Equilibrium values of output per-firm, profit-maximizing price, variety and tax revenue for an ad valorem and a unit tax under output equivalence are compared as follows:*

1. *ad valorem tax regime creates larger firms, i.e.,* $q_s < q_t$.
2. *Firms under ad valorem tax regime price lower, i.e.,* $p_s > p_t$.
3. *ad valorem tax creates less variety, i.e.,* $N_s > N_t$,
4. *The markup of firms is higher under unit tax regime.*
5. *Tax revenue and consumer surplus created by the ad valorem tax regime are* ***greater*** *compared to that of an output equivalent unit tax regime.*
6. *Welfare is* ***higher*** *under the ad valorem tax regime compared to that under an output equivalent unit tax regime.*

Proof Provided in "Appendix 1". □

Proposition one states that when firms homogenous in costs compete in a monopolistically competitive markets where consumers' love for variety is described by quadratic preferences which allow firms to adjust their markups, then the ad valorem tax is a welfare superior method of corrective tax. This result overturns the previous finding of welfare ranking with homogenous firms in a monopolistically competitive market by Dröge and Schröder (2009) where consumer preferences are characterized by Dixit–Stiglitz preferences that only allows for constant markups. Both models have two common forces at work. First, an ad valorem tax makes the demand curve more elastic. As a result, a profit-maximizing firm produces more at a lower price. This decreases the wedge between the price consumers pay and the price producers receive. Consumers therefore benefit from lower prices compared to a unit tax. Second, a unit tax acts like an increase in marginal cost. Firms therefore tend to decrease their output while they increase their prices. A higher operating surplus and smaller outputs are therefore sufficient to offset the fixed entry cost of entering the market. This allows more firms to exist in the market under unit taxes. More firms and hence more varieties under unit tax regime are appreciated by the consumers. However, lower prices are

too strong and dominates higher variety generated under unit tariff and overturns the consumer surplus ranking of Dröge and Schröder (2009).

In this model, furthermore, firms face identical residual demands under both regimes since we compare output equivalent tax rates. Output per-firm on the other hand is dictated by the free entry condition. In addition, as ad valorem tariff makes the demand curve more elastic, firms have a tendency to decrease their price (hence markup) and increase their outputs to maximize profits. If firms were forced to price under equal and constant markup pricing, firms performing under ad valorem tax regime would be constrained. Endogenous markup pricing allows firms to further decrease their prices to optimize their pricing behavior. This can be seen by the relative ranking of markups under both regimes. Thus, for a given unit tax rate, the ad valorem tax rate has to be more restrictive to generate equal market output. Hence, ad valorem tax regime generates higher tariff revenue.

4 Relative efficiency of tax regimes with heterogenous costs

I now reconstruct a model of heterogenous firms producing differentiated products. The marginal cost of production is ω. Potential entrants must pay a sunk entry cost f_e in order to learn their marginal cost of production. Once the entrants pay the entry cost, they draw their marginal costs ω from a common Pareto distribution with support on $[0, \overline{\omega}]$ and dispersion parameter k. Thus, marginal cost draws come from a cumulative distribution $G(\omega) = \left(\frac{\omega}{\overline{\omega}}\right)^k$, where parameter k indexes the dispersion of cost draws.

Once a firm is in the market, profit is maximized by taking the total number of varieties and average prices in the market as given. Since the entry cost f_e is sunk, firms that are able to cover their marginal costs survive in the market and continue to produce. This implies that there exists a marginal firm making zero profit (zero cutoff condition). The marginal firm's marginal cost of production is labeled as ω^*. Finally, prior to entry, the expected profit of a firm is $\int_0^{\omega^*} \pi(\omega) \mathrm{d}G(\omega) - f_e$. Firms will enter until this profit is driven to zero giving us the free entry condition. Notice that both the zero cutoff condition and the free entry condition can be written as functions of ω^*, number of firms and model parameters. Therefore, both cutoff marginal cost of production and number of firms are uniquely identified.

I further analyze the relative efficiency of ad valorem and unit taxes for a monopolistically competitive market of firms selling differentiated goods. I show that the heterogeneity of firms alters the superiority of ad valorem tax at higher levels of tax rate. Demand side and labor market for the homogenous good are identical with the model with symmetric costs. I further detail the market for differentiated goods.

Ad valorem tax regime

If the government collects t percent of price per unit of a good sold, then the profit-maximizing output of a firm with marginal cost of production ω is

$$q_t(\omega) = \frac{1}{\gamma}\left(p_t(\omega) - \frac{\omega}{1-t}\right). \tag{7}$$

Any potential entrant who pays the sunk entry cost and learns its marginal cost of production decides whether to stay in the market or exit immediately depending on the profit it can make. A marginal firm therefore makes zero profit. The expression below defines ω_t^* as the marginal cost of production for the marginal firm using Eq. (2).

$$\frac{\omega_t^*}{1-t} = \alpha - \eta Q_t. \tag{8}$$

The cutoff marginal cost ω_t^* summarizes the effects of both the average price and the number of firms on price and output of all firms. All of these measures can be written as functions of only the firm's own marginal cost ω and the marginal firm's ω_t^* as

$$\begin{aligned} q_t(\omega) &= \frac{1}{2\gamma}\frac{(\omega_t^* - \omega)}{1-t}, \\ p_t(\omega) &= \frac{1}{2}\frac{(\omega_t^* + \omega)}{1-t}. \end{aligned} \tag{9}$$

The free entry condition is also redefined as $\int_0^{\omega_t^*} \pi_t(\omega)\mathrm{d}G(\omega) - f_e$. The cutoff marginal cost is defined as functions of model parameters using output, price expressions [Eq. (9)] and free entry condition.

$$\omega_t^* = \left(f_e(k+1)(k+2)2\gamma\,(\overline{\omega})^k\,(1-t)\right)^{\frac{1}{k+2}}. \tag{10}$$

Total number of varieties N_t and total output Q_t are inferred from zero cutoff condition.

Unit tax regime

I now define output and price for all firms and the marginal firm under a unit tax policy where the government collects s for each unit sold. A firm's profit-maximizing output therefore becomes

$$q_s(\omega) = \frac{1}{\gamma}\left(p_s(\omega) - (\omega + s)\right). \tag{11}$$

The marginal cost of the marginal firm making zero profit in this economy is labeled as ω_s^*. Then, equating the profit of the marginal firm to zero gives us the zero cutoff condition as

$$\omega_s^* + s = \alpha - \eta Q_s. \tag{12}$$

I can rewrite output and price of all firms in terms of cutoff marginal cost, ω_s^*, and marginal cost of the firms, ω as

$$q_s(\omega) = \frac{1}{2\gamma}(\omega_s^* - \omega),$$
$$p_s(\omega) = \frac{1}{2}(\omega_s^* + \omega) + s. \tag{13}$$

Furthermore, the free entry condition defines the marginal firm by equating the ex-ante expected profits to the sunk entry cost paid. I can define ω_s^* as

$$\omega_s^* = \left(f_e(k+1)(k+2)2\gamma\,(\overline{\omega})^k \right)^{\frac{1}{k+2}}. \tag{14}$$

Notice that the cutoff marginal cost for the unit tax is independent of the tax rate. This is because a unit tax is fully passed on to consumers. Furthermore, all tax revenue collected by the government is redistributed back to consumers. The marginal firm is therefore not affected by the unit tax regime.

Proposition 2 *The equilibrium values of variety, tax revenue, consumer surplus, cut-off marginal cost of production for output equivalent ad valorem and unit taxes are compared as follows:*

1. *The marginal firm has lower marginal cost of production under an ad valorem tax regime, i.e.,* $\omega_s^* > \omega_t^*$. *Thus, the industry is more productive under ad valorem tax regime.*
2. *The variety created under an ad valorem tax is less than a unit tax, i.e.,* $N_t < N_s$.
3. *For a firm with marginal cost* ω, *markup difference under a unit tax regime and an output equivalent ad valorem tax regime is*

$$markup_s - markup_t = \frac{1}{2}\left(s - \frac{t\omega}{1-t}\right),$$

   which is positive for low levels of ω. *The difference in markups declines and eventually for* $\omega = \omega_t^*$ *the markup difference is negative.*
4. *The derivative of markup difference converges to* $-\infty$ *as* $t \longrightarrow 1$. *Also, as* $t \longrightarrow 0$, *the derivative is positive for low-cost firms and negative for high-cost firms.*
5. *The tax revenue created by the ad valorem tax regime is* ***less*** *compared to that of a unit tax regime,* $TR_s > TR_t$.
6. *The consumer surplus from the ad valorem tax regime is* ***greater*** *compared to that of a unit tax regime,* $CS_s < CS_t$.
7. *Welfare is* ***superior*** *under a unit tax regime compared to that of an ad valorem tax regime when unit taxes are higher.*

Proof Provided in "Appendix 1". □

I find that when cost asymmetry is added to a model of firms competing monopolistically and consumers having quadratic utility functions that allow firms to adjust their markups, a unit tax is welfare superior to an output equivalent ad valorem tax for higher values of tax requirements. This result challenges the parameter-independent superiority of ad valorem taxes with homogenous costs discussed in this paper. The

superiority of unit taxes is not only due to the appreciation of higher variety generated. There are two additional forces that are not present in models used to compare efficiency under different tax regimes. First, these tax instruments generate a different distribution of markups over marginal costs thus prices and outputs. Under ad valorem tax regime, relatively low-cost firms have lower markups and higher output compared to their markups and outputs under an output equivalent unit tax regime. On the other hand, relatively high-cost firms have higher markups and lower outputs again compared to their markups and outputs under a unit tax regime. This asymmetric distribution of markups increases the market shares of low-cost firms pricing relatively lower under ad valorem tax regime and hence results in lower tax revenue compared to an output equivalent unit tax regime. Although unit tax generates more varieties, relatively lower prices are appreciated by consumers more and as a result, consumer surplus is higher under ad valorem tax.

At low levels of tax requirements, total welfare is higher under ad valorem tax regime. Higher consumer surplus outweighs relatively lower tax revenue collected under ad valorem tax regime. As tax requirements increase, the response of firms under both regimes differ. First, the cutoff productivity of the marginal firm under unit tax is not a function of the tax rate since firms can fully pass the tax burden on consumers. On the other hand, when firms face an ad valorem tax regime, they pay part of the tax burden. Profit-maximizing low-cost firms, therefore, decrease their markups and increase their production in order to avoid paying this burden. This in return makes the market even more competitive and the marginal cost of the surviving marginal firm has to be even lower than that of the marginal firm under unit tax regime. This in return increases the concentration of market shares of low-cost firms. As a result, collected tax revenue declines more drastically under ad valorem tax regime. At higher tax requirements, the decline in tax revenue can not be compensated with higher consumer surplus (lower prices); hence, unit tax generates higher welfare. That is to say, heterogeneity exacerbates the price and variety gap between the two tax regimes, and variable markup pricing widens this gap even further.

Surprisingly, this model generates higher tax revenue under the unit tax regime. This might explain why most of the governments choose unit taxes as corrective tax instruments. However, an ad valorem tax makes the market more competitive and production concentrates at relatively low-cost firms. Thus, aggregate productivity is higher. On the other hand, the ability to pass the tax burden onto consumers under unit tax allows firms to survive at relatively lower output levels and hence production is spread out between more number of firms (varieties) at a wider range of productivity making the aggregate productivity of the industry lower.

Schröder and Sørensen (2010) analyze equal yield unit and ad valorem tax regimes in a model of monopolistic competition with heterogenous firms and Dixit–Stiglitz preferences. They find that ad valorem tax regime is superior. When I replace Dixit–Stiglitz preferences with quadratic preferences, with a numerical example, I show that unit tax generates higher welfare than an equal yield ad valorem tax at higher revenue requirements.[6]

[6] A numerical example is presented in "Appendix 2".

5 Conclusion

The effect of market structure and cost asymmetries of unit and ad valorem taxes on welfare of consumers has been extensively discussed. This paper studies the welfare implications of output equivalent unit and ad valorem taxes in a model of monopolistic competition with heterogenous firms and with a utility function that allows for variable markups. The discussion contributes mainly to the public economics and the environmental economics literature on corrective taxes where little has been done with heterogenous firms and specifically variable markup pricing.

Without cost asymmetries, an ad valorem tax is superior. However, when including cost asymmetries, I show that a unit tax dominates an ad valorem tax for higher levels of tax requirements. This parameter-dependent ranking is a result of both heterogeneity and variable markup pricing. Both effects increase the gap between the price and variety generated under these two tax regimes. Furthermore, as ad valorem tax rate increases, the market under an ad valorem tax becomes so competitive that production is concentrated among relatively low-cost firms that has relatively lower markups compared to their counterparts under unit tax regime. As a result, the difference between the generated tax revenue increases in favor of unit tax such that at high tax requirements, higher tax revenue outweighs ad valorem tax regimes ability to generate higher consumer surplus.

Compliance with ethical standards

Conflict of interest The author acknowledges that she receives no financial support. The author also declares that she has no conflict of interest.

Human and animal rights This article does not contain any studies with human or animals participants performed by the author.

Informed consent This article does not contain any information that requires informed consent or does not use any data (private–public).

Appendix 1

Proof of Proposition 1 (Symmetric cost case) First, I find the unit tax rate that is the total output equivalent to a given ad valorem tax rate, t. I rearrange inverse demand functions [Eq. (2)] for two tax regimes. I find that $\alpha - \eta Q_t = p_t + \gamma q_t$, and similarly, $\alpha - \eta Q_s = p_s + \gamma q_s$. Since our comparison unit is equal to total output, I conclude that

$$p_t + \gamma q_t = p_s + \gamma q_s. \tag{15}$$

When I use the supply function [Eqs. (3) and (5)] and equilibrium firm outputs and prices for both regimes [Eqs. (4) and (6)], I find the output equivalent unit tax rate as a function of model parameters:

$$s = \frac{tc}{1-t} + 2\sqrt{\gamma f_e}\left(\frac{1}{\sqrt{1-t}} - 1\right). \tag{16}$$

Firms under a unit tax regime are smaller and prices are higher than those under an ad valorem tax regime. The output per-firm comparison is straightforward from the equilibrium values, $q_t = \sqrt{\frac{f_e}{\gamma(1-t)}}$ and $q_s = \sqrt{\frac{f_e}{\gamma}}$. For any $t \in (0,1)$, $q_s < q_t$. Furthermore, following Eq. (15), I can conclude that $p_s > p_t$. Similarly, the total output equivalence gives fewer total varieties under an ad valorem tax compared to a unit tax regime; $N_s > N_t$. The difference between the markups over marginal cost under both tax regime is $\frac{p_t}{c} - \frac{p_s}{c} = \frac{\gamma q_s - \gamma q_t}{c} = \frac{\sqrt{\gamma f_e}}{c}(1 - \frac{1}{\sqrt{1-t}}) < 0$.

Tax revenue on the other hand is not straightforward. Since the total output level is equal, one can simply compare the tax revenue collected per unit of output. Under a unit tax, this value is s, whereas under an ad valorem tax regime, the per unit tax revenue is tp_t. From Eq. (16), $s = \frac{tc}{1-t} + 2\gamma(q_t - q_s)$. Similarly, $tp_t = t\gamma q_t + \frac{tc}{1-t}$. The difference between the tax revenues per unit is therefore equal to

$$\begin{aligned}\frac{\mathrm{TR}_t - \mathrm{TR}_s}{Q} &= tp_t - s \\ &= \gamma\,(2q_s - (2-t)q_t) \\ &> 0 \quad \text{for all } t \in (0,1).\end{aligned}$$

Consumer surplus under both regimes is utility created by consuming both the homogenous good and differentiated goods net of money spent on them. Consumer surplus under both regimes can therefore be written as

$$\begin{aligned}\mathrm{CS}_t &= q_o + \alpha Q_t - \frac{1}{2}\gamma N_t (q_t)^2 - \frac{1}{2}\eta (Q_t)^2 - p_t Q_t - p_o q_o \\ \mathrm{CS}_s &= q_o + \alpha Q_s - \frac{1}{2}\gamma N_s (q_s)^2 - \frac{1}{2}\eta (Q_s)^2 - p_s Q_s - p_o q_o.\end{aligned}$$

Tax regimes are output equivalent and price of the homogenous good is normalized to 1. If I use these two pieces of information and take the difference between the two consumer surpluses, I find

$$\begin{aligned}\mathrm{CS}_s - \mathrm{CS}_t &= Q\left(\frac{1}{2}\gamma q_t - \frac{1}{2}\gamma q_s + \gamma q_s - \gamma q_t\right) \\ &= Q\left(\frac{1}{2}\gamma q_s - \frac{1}{2}\gamma q_t\right).\end{aligned}$$

Using Eq. (15), I can further simplify the above expression to $\mathrm{CS}_s - \mathrm{CS}_t = Q(\frac{1}{2}\gamma q_s - \frac{1}{2}\gamma q_t)$. Since $q_s < q_t$, the difference in consumer surpluses under two regimes is

$CS_s - CS_t < 0$ for all model parameters and $t \in (0, 1)$. Finally, welfare under an ad valorem tax regime is higher compared to an output equivalent unit tax regime because it creates greater consumer surplus and tax revenue. □

Proof of Proposition 2 (Asymmetric cost case) I first compare cutoff marginal cost values under both regimes. Equations (10) and (14) give the direct relationship between ω_s^* and ω_t^* as $\omega_t^* = (1-t)^{\frac{1}{k+2}}\omega_s^*$. Since t is between 0 and 1, I conclude that $\omega_t^* < \omega_s^*$.

I use equal output criterion with redistributed tax revenue in order to show the relative ranking of the variety created. The output created by all the firms in the market under an ad valorem tax is given as

$$\begin{aligned} Q_t &= \int_0^{\omega_t^*} q(w)\mathrm{d}G(w|w < \omega_t^*) \\ &= \int_0^{\omega_t^*} \frac{1}{2\gamma}(\omega_t^* - \omega)\mathrm{d}G(w|w < \omega_t^*) \\ &= \frac{1}{2\gamma(k+1)}\frac{\omega_t^*}{1-t}N_t. \end{aligned}$$

Similarly, the total output under a unit tax regime is $Q_s = \frac{1}{2\gamma(k+1)}\omega_s^* N_s$, and additionally, $\frac{\omega_t^*}{1-t} > \omega_s^*$. Furthermore, I defined the equivalency of these two regimes through output; as a result, $N_s > N_t$. A unit tax creates greater variety than the output equivalent ad valorem tax.

The difference in markups for a given firm under two tax regimes is $\text{markup}_s - \text{markup}_t = \frac{1}{2}(s - \frac{t\omega}{1-t})$. If I write the expression in terms of model parameters and ad valorem tax rate, I get $\frac{1}{2}\left[\omega_s^*\left((1-t)^{-(k+1)/(k+2)} - 1\right) - \frac{t\omega}{1-t}\right]$. The expression continuously decreases for ω. Also the expression is positive for very low levels of ω. Furthermore, at $\omega = \omega_t^*$, the difference is negative. $\frac{1}{2}(s - \frac{t\omega}{1-t}) = \frac{1}{2}(\frac{\omega_t^*}{1-t} - \omega_s^* - \frac{t\omega_t^*}{1-t}) = \frac{1}{2}(\omega_t^* - \omega_s^*) < 0$.

The derivative of the markup difference with respect to t is:

$$\frac{1}{2}\omega_s^*\frac{k+1}{k+2}(1-t)^{-1-((k+1)/(k+2))} - \frac{\omega}{(1-t)^2}.$$

The derivative goes to $-\infty$ as $t \to 1$. The expression goes to $\frac{1}{2}\omega_s^*\frac{k+1}{k+2} - \omega$ as $t \to 0$. For very low tax requirements, the difference in markups increases at low-cost levels. This implies that markup of a firm under ad valorem tax rate decreases further for low-cost firms. The expression is negative for high-cost firms. This implies that the markup difference declines for high-cost firms at lower tax requirements.

Total tax collected from a firm with a marginal cost of production ω is $tp_t(w)q_t(w)$. In order to calculate the total tax collected from all firms, I aggregate per-firm collected tax over their respective marginal cost values given as

$$\begin{aligned}\mathrm{TR}_t &= \int_0^{\omega_t^*} tp(w)q(w)\mathrm{d}G(w|w<\omega_t^*)\\ &= \int_0^{\omega_t^*} t\frac{1}{4\gamma}\frac{(\omega_t^*+\omega)}{1-t}\frac{(\omega_t^*-\omega)}{1-t}\mathrm{d}G(w|w<\omega_t^*)\\ &= \frac{1}{2\gamma}\frac{1}{k+2}\frac{t}{(1-t)^2}(\omega_t^*)^2N_t. \end{aligned} \tag{17}$$

If I further simplify the expression using the total output Q_t calculated above, I find

$$\mathrm{TR}_t = \frac{t}{1-t}\frac{k+1}{k+2}\omega_t^* Q_t. \tag{18}$$

I similarly aggregate tax collected from a single firm with a marginal cost of production ω in order to calculate the total tax revenue created under a unit tax. The per-firm tax revenue is $sq_s(w)$. If I aggregate over all surviving firms in the market, I find that TR_s equals $\frac{1}{2\gamma}\frac{1}{k+1}s\omega_s^*N_s$ or simply sQ_s. In order to compare the tax revenue collected under both regimes, I take the difference in tax revenues calculated above:

$$\mathrm{TR}_s - \mathrm{TR}_t = sQ_s - \frac{t}{1-t}\frac{k+1}{k+2}\omega_t^* Q_t.$$

Since the total output is equal under both regimes, I use the total output as a common factor. Furthermore, by using the zero cutoff conditions in Eqs. (8), (12), and the equivalency definition, I find the output equivalent tax rate s as a function of a given ad valorem rate t as $s = \frac{\omega_t^*}{1-t} - \omega_s^*$. If I insert in the expression for s and common factor Q, I find

$$\mathrm{TR}_s - \mathrm{TR}_t = Q\left(\frac{\omega_t^*}{1-t} - \omega_s^* - \frac{t}{1-t}\frac{k+1}{k+2}\omega_t^*\right).$$

If I insert the equality between the cutoff rates driven from free entry condition $\omega_t^* = (1-t)^{\frac{1}{k+2}}\omega_s^*$ back into the above expression, the expression becomes a function of t, k and common multipliers:

$$\mathrm{TR}_s - \mathrm{TR}_t = Q\omega_s^*\left(\frac{(1-t)^{\frac{1}{k+2}}}{1-t} - 1 - \frac{t}{1-t}\frac{k+1}{k+2}(1-t)^{\frac{1}{k+2}}\right).$$

The expression above is a continuous function for $t \in (0,1)$, and at $t = 0$, the difference is zero. Furthermore, $\lim_{t\to 1}(\mathrm{TR}_s - \mathrm{TR}_t) = \infty$. Finally, $\frac{\mathrm{d}(\mathrm{TR}_s - \mathrm{TR}_t)}{\mathrm{d}t} = \frac{k+1}{(k+2)^2}(1-t)^{-(k+1)/(k+2)}\frac{t}{1-t} > 0$ for all t values between 0 and 1. This means that the difference in tax revenues under a unit tax and an output equivalent ad valorem tax is not only always positive but is also increasing for $t \in (0,1)$.

The consumer surplus under a tax regime is the utility driven by consuming the differentiated products and the homogenous good minus the cost of consuming those goods. The consumer surplus under an ad valorem tax is therefore

$$\begin{aligned}\mathrm{CS}_t &= q_0 + \alpha Q_t - \frac{1}{2}\gamma \int_0^{\omega_t^*} q_t^2(w)\mathrm{d}G(w|w < \omega_t^*) \\ &\quad - \frac{1}{2}\eta(Q_t)^2 - p_0 q_0 - \int_0^{\omega_t^*} p_t(\omega) q_t(w)\mathrm{d}G(w|w < \omega_t^*) \\ &= q_0 + \alpha Q_t - \frac{1}{2(k+2)}\frac{\omega_t^*}{1-t}Q_t - \frac{1}{2}\eta(Q_t)^2 - p_0 q_0 - \frac{k+1}{k+2}\frac{\omega_t^*}{1-t}Q_t. \quad (19)\end{aligned}$$

Similarly, the consumer surplus under a unit tax is

$$q_0 + \alpha Q_s - \frac{1}{2(k+2)}\omega_s^* Q_s - \frac{1}{2}\eta(Q_s)^2 - p_0 q_0 - \frac{k+1}{k+2}\omega_s^* Q_s - s Q_s. \quad (20)$$

If I calculate the difference between the consumer surplus under a unit tax and an ad valorem tax taking into account that the price of a homogenous good is one and the total output levels are equal at Q, I find

$$\begin{aligned}\mathrm{CS}_s - \mathrm{CS}_t &= -\frac{1}{2(k+2)}\omega_s^* Q - \frac{k+1}{k+2}\omega_s^* Q - sQ + \frac{1}{2(k+2)}\frac{\omega_t^*}{1-t}Q + \frac{k+1}{k+2}\frac{\omega_t^*}{1-t}Q, \text{ and} \\ &= Q\left(-\frac{1}{2(k+2)}\omega_s^* - \frac{k+1}{k+2}\omega_s^* - s + \frac{1}{2(k+2)}\frac{\omega_t^*}{1-t} + \frac{k+1}{k+2}\frac{\omega_t^*}{1-t}\right).\end{aligned}$$

I insert unit tax rate s as a function of the cutoff marginal cost values $s = \frac{\omega_t^*}{1-t} - \omega_s^*$ back into the above expression and find

$$\begin{aligned}\mathrm{CS}_s - \mathrm{CS}_t &= Q\frac{1}{2(k+2)}\left(\omega_s^* - \frac{\omega_t^*}{1-t}\right) \\ &< 0.\end{aligned}$$

Thus, the consumer surplus generated under an ad valorem tax is always greater than the consumer surplus generated under a unit tax.

Welfare on the other hand does not have a clear ranking. Under a unit tax regime, welfare is therefore

$$\begin{aligned}W_s &= \mathrm{CS}_s + \mathrm{TR}_s \\ &= q_0 + \alpha Q_s - \frac{1}{2(k+2)}\omega_s^* Q_s - \frac{1}{2}\eta(Q_s)^2 - p_0 q_0 \\ &\quad - \frac{k+1}{k+2}\omega_s^* Q_s - sQ_s + sQ_s\end{aligned}$$

$$= Q_s \left(\alpha - \frac{1}{2(k+2)} \omega_s^* - \frac{1}{2} \eta Q_s - \frac{k+1}{k+2} \omega_s^* \right).$$

Similarly, welfare under an ad valorem tax is

$$\begin{aligned} W_t &= \mathrm{CS}_t + \mathrm{TR}_t \\ &= q_0 + \alpha Q_t - \frac{1}{2(k+2)} \frac{\omega_t^*}{1-t} Q_t - \frac{1}{2} \eta (Q_t)^2 - p_0 q_0 - \frac{k+1}{k+2} \frac{\omega_t^*}{1-t} Q_t \\ &\quad + \frac{t}{1-t} \frac{k+1}{k+2} \omega_t^* Q_t \\ &= Q \left(\alpha - \frac{1}{2(k+2)} \frac{\omega_t^*}{1-t} - \frac{1}{2} \eta Q_t - \frac{k+1}{k+2} \frac{\omega_t^*}{1-t} + \frac{t}{1-t} \frac{k+1}{k+2} \omega_t^* \right). \end{aligned}$$

By taking the difference between both regimes, I find

$$\begin{aligned} W_s - W_t &= Q \left(-\frac{1}{2(k+2)} \omega_s^* - \frac{k+1}{k+2} \omega_s^* + \frac{1}{2(k+2)} \frac{\omega_t^*}{1-t} \right. \\ &\quad \left. + \frac{k+1}{k+2} \frac{\omega_t^*}{1-t} - \frac{t}{1-t} \frac{k+1}{k+2} \omega_t^* \right) \\ &= Q \omega_s^* \frac{1}{k+2} \left(-k - \frac{3}{2} + \frac{1}{2} \frac{(1-t)^{1/(k+2)}}{(1-t)} + (k+1)(1-t)^{1/(k+2)} \right). \end{aligned}$$

The expression above is zero for $t = 0$ and is continuous for $t \in (0, 1)$. Furthermore, $\lim_{t \to 1} (W_s - W_t) = \infty$. The derivative of the difference function with respect to t is $Q \omega_s^* \frac{k+1}{(k+2)^2} (1-t)^{-(k+1)/(k+2)} \frac{2t-1}{1-t}$. The derivative is positive for $t > 0.5$ and zero for $t = 0.5$. This implies that the difference function begins at zero and decreases up to $t = 0.5$. The difference reaches its minimum at $t = 0.5$ and begins to increase until it reaches ∞ for $t = 1$. Thus, welfare under a unit tax regime is greater than an output equivalent ad valorem tax regime if t is large. □

Appendix 2

I challenge the result of Schröder and Sørensen (2010) with a numerical example. The algorithm for the solution is explained in the following steps as given:

Step 1: For any given ad valorem tax rate t, calculate the marginal cost of the marginal firm under both regimes using model parameters and Eqs. (10) and (14).

Step 2: Use the zero cutoff conditions for both regimes [Eqs. (8) and (12)] and find the formulas for total production Q_s and Q_t. Plug these expressions back into the revenue functions [Eqs. (17) and (18)].

Step 3: Solve for a unit tax rate that gives us equal tax revenue as the tax revenue collected under ad valorem tax regime with tax rate, t.

Step 4: For given t and numerically solved equal yield unit tax rate s, calculate consumer surplus under both regimes using Eqs. (19) and (20).

Table 1 Parameter values

Model parameters	α	γ	η	k	L	fe	$\overline{\omega}$
Assignments	69.213	3	900	2	1	0.5	7

Step 5: Update the ad valorem tax rate and go to step 1.

There are also model specific parameter restrictions. The first restriction is on the upper bound of the marginal cost distribution, $\overline{\omega}$. $\overline{\omega}$ has to be greater that $\sqrt{2(k+1)(k+2)\gamma f_e/L}$. This guarantees that marginal cost of the marginal firm under unit tax is less that the upper bound on the marginal cost of production. The remaining two restrictions are on α. α parametrizes the relative demand for differentiated goods. It has to be high enough so that there is a positive demand for differentiated goods, i.e., $P_{\max} = \omega_s^* + s < \alpha$. Also α cannot be so high because demand for the numeraire good has to be positive, hence $\alpha < 2\sqrt{\eta(k+1)/(k+2)}$. The set of parameters listed in Table 1 satisfies all parameter restrictions and the following figures are drawn using these parameters and following the algorithm explained above.

Figure 1 demonstrates the difference in consumer surplus under ad valorem tax regime and unit tax regime. The tax rate, t, is on the horizontal axis. The numerical example shows that welfare under unit tax is higher for high levels of tax requirements. The reason for this parameter-dependent result is due to the excessive exit of firms under ad valorem tax regime at higher tax rates. The difference between the number of firms and hence varieties is drawn in Fig. 2. As the tax rate increases, the varieties under ad valorem tax regime decreases and the difference between unit and ad valorem tax regime converges to a relatively high negative number.

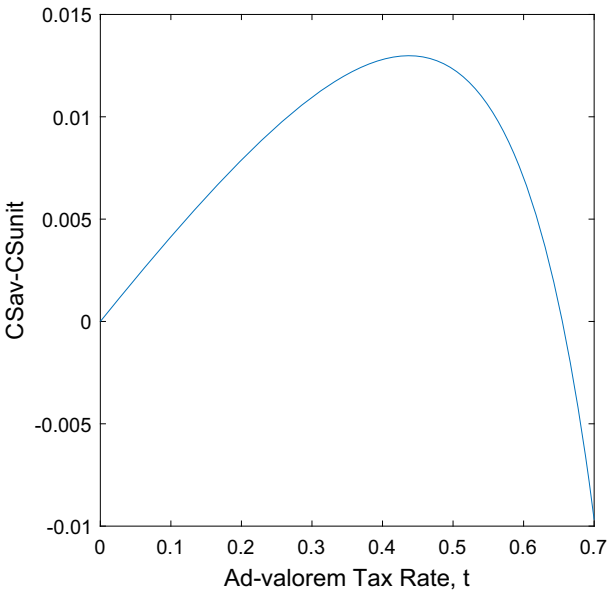

Fig. 1 Difference in consumer surplus

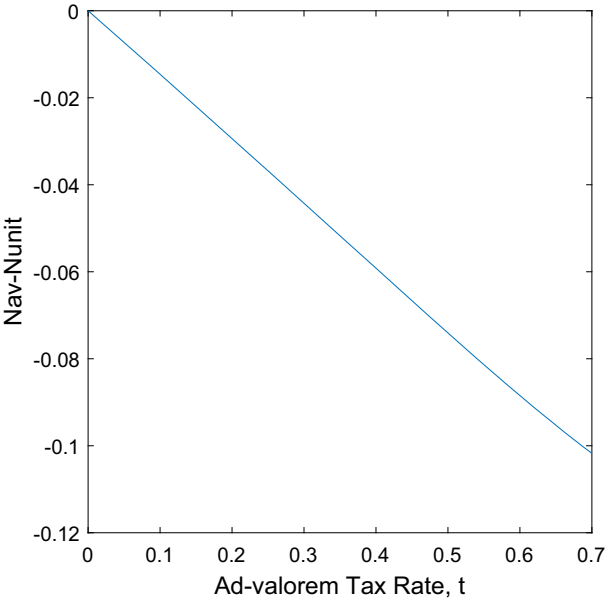

Fig. 2 Difference in varieties

References

Arkolakis C, Costinot A, Donaldson D, Rodriguez-Clare A (2019) The elusive pro-competitive effects of trade. Rev Econ Stud (forthcoming)

De Loecker J, Eeckhout J, Unger G (2019) The rise of market power and the macroeconomic implications. Working paper

Dröge S, Schröder PJH (2009) The welfare comparison of corrective ad valorem and unit taxes under monopolistic competition. Int J Tax Public Finance 16:164–175

Edmond C, Midrigan V, Yu DX (2015) Competition, mark-ups, and gains from international trade. Am Econ Rev 105(10):3183–3221

Keen M (1998) The balance between specific and ad-valorem taxation. Fisc Stud 19:1–37

Melitz M, Ottaviano G (2008) Market size, trade, and productivity. Rev Econ Stud 75:295–316

Pirttilä J (2002) Specific versus ad valorem taxation and externalities. J Econ 76:177–187

Schröder PJH (2004) The comparison between ad-valorem and unit taxes under monopolistic competition. J Econ 83:281–292

Schröder PJH, Sørensen A (2010) Ad valorem versus unit taxes: monopolistic competition, heterogeneous firms, and intra-industry reallocations. J Econ 101(3):247–265

Seely A (2009) Landfill tax: introduction and early history. House of Commons, Library, Business and Transportation section, Standard note: SN/BT/237

Skeath S, Trendel AG (1994) A Pareto comparison of ad valorem and unit taxes in noncompetitive environments. J Public Econ 53(1):53–71

Suits DB, Musgrave RA (1953) Ad valorem and unit taxes compared. Q J Econ 67:598–604