Online Appendix
Not Intended for Journal Publication

"VAT Notches, Voluntary Registration and Bunching:
Theory and UK Evidence"

September 2019

Li Liu (International Monetary Fund and Oxford CBT)
Ben Lockwood (University of Warwick, Oxford CBT, CEPR)
Miguel Almunia (CUNEF, CAGE, CEPR and Oxford CBT)
Eddy H. F. Tam (Oxford CBT)

A Theoretical Appendix

A.1 Proof of Propositions 1 and 2 in the Paper

Proof of Proposition 1. We will assume $e_C = e_B = e$. (a) If registered, the firm of type a chooses p_C, p_B to maximize (11) subject to (12),(13) when $I = 1$. This problem is easily solved to give prices as a mark-up over cost i.e.

$$p_C(a) = p_B(a) = p(a) = \frac{e}{e-1}c(1;a). \tag{A.1}$$

Substituting (A.1) back into the profit function (11), and using $e_C = e_B = e$, we can derive the following formula for maximized profit:

$$\pi(1;a) = \kappa(\lambda(1+t)^{-e} + (1-\lambda)A_B)c(1;a)^{1-e}, \;\kappa = \left(\frac{e}{1-e}\right)^{-e}\frac{1}{1-e} \tag{A.2}$$

Now consider voluntary registration. A necessary and sufficient condition for this is that profit with registration is greater than profit without, ignoring the constraint that sales be below the threshold. So, we consider the problem where the firm chooses p_C, p_B to maximize (11) subject to (12),(13) when $I = 0$, ignoring the sales constraint. Solving this problem, following the steps for the registered firm, we get maximized profit of

$$\pi(0;a) = \kappa(\lambda + (1-\lambda)A_B)c(0;a)^{1-e} \tag{A.3}$$

Then, the voluntary registration condition is $\pi(1;a) \geq \pi(0;a)$. After some simple rearrangement, using (A.2),(A.3), this reduces to (16).

(b) We only need show that T in (16) is increasing in λ and decreasing in ω. The statement for λ is obvious by inspection of (15). For ω, note first that $1 + \Delta_c \equiv 1 + \omega t$ is increasing in ω. So, $(1+\Delta)^{1-e}$ is decreasing in ω as $e > 1$. So, T is decreasing in ω as required. □

Proof of Proposition 2. We first prove part (a). (i) First, we characterize the relationship between firm type a and sales s, under the assumption that the firm does not have to register for VAT. From (12) and (A.1), the optimal price and total sales (B2C and B2B) of a non-registered firm are;

$$p(a) = \frac{e}{e-1}\frac{1+\omega t}{a}, \; s(a) = (\lambda + (1-\lambda)A_B)(p(a))^{1-e} \tag{A.4}$$

Combining these two expressions to substitute out $p(a)$, we see that the relationship between firms type and sales is:

$$s = (\lambda + (1-\lambda)A_B)\left(\frac{e(1+\Delta_c)}{e-1}\right)^{1-e} a^{e-1} \tag{A.5}$$

(ii) We now write the indifference condition (17) that determines $a^* + \Delta a^*$ explicitly in the space of sales. To To lighten notation, and using $e_C = e_B = e$, define

$$A_0 = \lambda + (1-\lambda)A_B, \;\; A_1 = \lambda(1+t)^{-e} + (1-\lambda)A_B$$

First, from (A.2), and using the restriction that $e_C = e_B = e$, a firm that registers has maximized profit:

$$\pi(1;a) = A_1 \left(\frac{e}{1-e}\right)^{-e} \frac{a^{e-1}}{1-e} \tag{A.6}$$

Next, the profit from being just at the VAT threshold for an $a-$type when constrained is

$$\pi(0;a) = s^* - \frac{1+\Delta_c}{a}\frac{s^*}{p} \tag{A.7}$$

But solving for p from the constraint $A_0 p^{1-e} = s^*$, we get $p = \left(\frac{s^*}{A_0}\right)^{-1/(e-1)}$. Substituting this back into (A.7), and using $a = a^* + \Delta a^*$, we get

$$\pi(0;a^* + \Delta a^*) = s^* - \frac{1+\Delta_c}{A_0^{1/(e-1)}(a^* + \Delta a^*)}(s^*)^{e/(e-1)} \tag{A.8}$$

Then combining (A.5), evaluated at $a = a^* + \Delta a^*$, and (A.8), we get:

$$\begin{aligned} \pi(0;a^* + \Delta a^*) &= s^* - \frac{1+\Delta_c}{A_0^{1/(e-1)}(a^* + \Delta a^*)}(s^*)^{e/(e-1)} \\ &= s^* - (s^*)^{e/(e-1)}\left(\frac{e-1}{e}\right)(s^* + \Delta s^*)^{1/(1-e)} \end{aligned} \tag{A.9}$$

Again using (A.5), evaluated at $a = a^* + \Delta a^*$, in (A.6), we get:

$$\begin{aligned}\pi(1; a^* + \Delta a^*) &= A_1 \left(\frac{e}{1-e}\right)^{-e} \frac{1}{1-e}(a^* + \Delta a^*)^{e-1} \\ &= \frac{A_1}{A_0}\left(\frac{e}{1-e}\right)^{-e} \frac{1}{1-e}\left(\frac{e\,(1+\Delta_c)}{e-1}\right)^{e-1}(s^* + \Delta s^*) \\ &= \frac{T}{e}(s^* + \Delta s^*)\end{aligned} \tag{A.10}$$

So, using (A.10),(A.9), the indifference condition $\pi(1; a^* + \Delta a^*) = \pi_0(0; a^* + \Delta a^*)$ becomes

$$s^* - (s^*)^{e/(e-1)}\left(\frac{e-1}{e}\right)(s^* + \Delta s^*)^{1/(1-e)} - \frac{T}{e}(s^* + \Delta s^*) = 0 \tag{A.11}$$

After some simplification of (A.11) (divide through by s^*, then $1 + \frac{\Delta s^*}{s^*}$, and multiply by e) we get (18) as required.

To prove part (b), first, (18) can be rewritten as

$$f(x, e) - T(\lambda, \omega, e) = 0, f(x) \equiv ex - (e-1)x^{e/(e-1)}, x = \frac{1}{(1 + \Delta s * /s*)} \tag{A.12}$$

So, from (A.12):

$$\frac{dx}{d\lambda} = \frac{T_\lambda}{f_x}, \ \frac{dx}{d\omega} = \frac{T_\omega}{f_x}, \ \frac{dx}{de} = \frac{T_e - f_e}{f_x} \tag{A.13}$$

Moreover, note that

$$f_x = e(1 - x^{1/(e-1)}) > 0 \tag{A.14}$$

because $x < 1$, and $e > 1$, so $x^{1/(e-1)} < 1$. Also, we know that $T_\lambda < 0$, $T_\omega > 0$ and so from (A.13), (A.14), we conclude that $\frac{dx}{d\lambda} < 0$, $\frac{dx}{d\omega} > 0$. As x is an inverse measure of bunching, it follows that as λ increases, Δs^* rises, and ω rises, Δs^* falls.

Finally, from (A.13),

$$f_e = x - x^{e/(e-1)} - (e-1)\frac{d(x^{e/(e-1)})}{de} = x - x^{e/(e-1)} - \frac{\ln x}{e-1}x^{e/(e-1)}$$

So, for $f_e > 0$, we require

$$x^{-1/(e-1)} > 1 + \frac{\ln x}{e-1}$$

But as $x < 1$, $\ln x < 0$, so we require $x^{1/(e-1)} < 1$, which certainly holds as $x < 1$ and $1/(e-1) > 0$. So, we conclude that $f_e > 0$. It then follows from (A.13) ,(A.14), that $\frac{dx}{de} < 0$ as long as $T_e < 0$. But then bunching increases if $T_e < 0$, as claimed. □

A.2 Other Results

Proof that Voluntary Registration is Not Possible in the de Paula and Scheinkman (2010) Model

We follow the notation of their paper. There are two kinds of firms, upstream firms, and downstream firms. An informal downstream firm is by definition, not registered for VAT, and get gets profit $\pi_N = \max_{x \leq \overline{x}}\{\theta^d x^\alpha - p_i x\}$, where x is sales, $\overline{x}$ is the registration threshold, and p_i is the price of the input of bought from an upstream informal firm. If this firm registers, it gets profit $\pi_R = \max_x\{(1-\tau)(\theta^d x^\alpha - p_f x)\}$, where τ is the rate of VAT, and $p_f > p_i$ is the cost of buying from the formal sector. Voluntary registration involves registering at a turnover (say $x_0 < \overline{x}$), giving profit $(1-\tau)(\theta^d x_0^\alpha - p_f x_0)$. But then, as $\tau > 0,\ p_f > p_i$;

$$\pi_R = (1-\tau)(\theta^d x_0^\alpha - p_f x_0) < \theta^d x_0^\alpha - p_f x_0 < \theta^d x_0^\alpha - p_i x_0 \leq \pi_N$$

so that the downstream firm can do better by not registering. Now consider an upstream firm in their model. This firm gets $\pi_N = \min \theta_u, \overline{x}$ if not registered, and $\pi_R = (1-\tau)\theta_u$ if registered. So, such a firm registers voluntarily iff $\theta_u < \overline{x}$. But, for such a firm, $\pi_N = \theta_u > (1-\tau)\theta_u = \pi_R$; again, the downstream firm can do better by not registering.

General Proof that Voluntary Registration is not Possible with Only B2C Sales

Consider a firm facing a residual demand function from final consumers of $x(q)$, where q is the consumer price. This covers both the cases of monopoly, where $x(.)$ is also the actual demand curve, and monopolistic competition, where $x(.)$ is demand for that firm's product, taking the prices of all other firms as fixed. Assume all sales are to final consumers i.e. B2C. If a firm is registered for VAT, profit is then

$$\pi_R(p) = px(p(1+t)) - c(x(p(1+t)), w, r)$$

where p is the producer price, and $c(x,\ w, r)$ is the cost function given output x, and prices of labour and the intermediate input w, r. This is completely general cost function that clearly includes the cost function in the paper as a special case. If the firm is not registered for VAT, profit is

$$\pi_N(p) = px(p) - c(x(p), w, r(1+t))$$

Then, we have

$$\begin{aligned}
\pi_R &= \max_p \{px(p(1+t)) - c(x(p(1+t)), w, r)\} \\
&= \max_q \left\{ \frac{q}{1+t} x(q) - c(x(q), w, r) \right\} \\
&= \frac{1}{1+t} \max_q \{qx(q) - (1+t)\, c(x(q), w, r)\} \\
&< \max_q \{(q - c(x(q), w, r(1+t)))\} \\
&= \pi_N
\end{aligned}$$

So, with only B2C sales, no firm would every wish to register voluntarily.

A.3 The Effect of e on T

Table A.1 shows the value of T for as e increases, for different combinations of λ and ω.

Table A.1. The Effect of Competition on the Sufficient Statistic T

e	2	5	10	20	50	100
$\omega = 0.9, \lambda = 0.1$	1.039	1.408	1.967	2.361	1.564	0.678
$\omega = 0.9, \lambda = 0.9$	0.826	0.796	0.739	0.629	0.381	0.164
$\omega = 0.1, \lambda = 0.1$	0.898	0.786	0.530	0.148	0.001	0.000
$\omega = 0.1, \lambda = 0.9$	0.714	0.444	0.199	0.039	0.000	0.000

The interpretation of these results is as follows. In the first row, the input cost ratio is very high, and the share of B2C sales is very low. These are the conditions under which we expect voluntary registration ($T > 1$) to occur, and indeed this is the case. Note here that T is not monotonic in e; it first rises and then falls. This non-monotonicity is due to the fact that for low values of e, the effect of increasing competition on the input VAT component of T can dominate the effect on the output component of T. The second row is where both the input cost ratio and the share of B2C sales are very high. Here, these two opposing forces always lead to a T below 1 that is monotonically decreasing in T. The same qualitative picture as in the second row also emerges when the input cost ratio is low, for both a low and high B2C ratio.

B Extensions to VAT Evasion and Compliance Costs

Here, we model the simplest and most common form of VAT evasion, under-reporting of sales. This has two aspects. First, a non-registered firm can hide a share of sales, for example by using cash transactions. It is widely believed that the VAT chain makes it more difficult to conceal sales to other registered businesses, so we will assume that only that a share ν_N of B2C sales can be hidden. Then, sales can be s high as $s^* + \nu_N px$ without registering for VAT.

We also allow registered sellers can use this strategy, albeit at a higher cost of detection. So, suppose that such a seller does not charge VAT on some proportion ν_R of B2C sales px. The total cost of x units of the good to the household, if purchased from a registered seller, will be

$$\nu_R xp + (1 - \nu_R)xp(1 + t) = xp(1 + (1 - \nu_R)t).$$

That is, the household faces an average price of $p(1 + (1 - \nu_R)t)$, and the firm continues to get revenue p on every unit sold to the household.

We will assume that ν_N, ν_R are exogenously fixed, both for simplicity, and also because there are some analytical issues in endogenizing them.[27] The main qualitative points will extend to the endogenous case.[28] We will assume that $0 \leq \nu_R \leq \nu_N < 1$, reflecting the fact that non-registered firms are less likely to be audited than registered ones.

It is then easily verified that for voluntary registration, the analysis proceeds much as before except that the VAT sufficient statistic becomes

$$T(\nu) = (1 - \Delta_d(\nu_R))(1 + \Delta_c)^{e-1}, \ \Delta_d(\nu_R) = \frac{\lambda(1 - (1 + (1 - \nu_R)t)^{-e})}{\lambda + (1 - \lambda)A_B} \tag{B.15}$$

Thus, with evasion, the output VAT effect depends on ν_R and is smaller than without evasion

[27] For example, suppose that the registered firm chooses ν_R to maximize profit minus evasion cost $g(\nu_R)$. It is easily verified that optimized profit is convex in ν_R, as it only depends on ν_R via the term $(1 + t(1 - \nu_R))^{-e}$, which is a convex function of ν_R. So, to have an interior solution, $g(.)$ also has to be sufficiently convex in ν_R. But then, a closed-form solution for ν_R cannot be found.

[28] For example, suppose the cost of evasion is linear in ν, up to a limit $\overline{\nu} < 1$. Then, as profit is convex in ν, as explained in the previous footnote, and the evasion cost is small, the firm will always choose $\overline{\nu}$, so that it is effectively exogenous.

i.e. $\Delta_d(\nu_R) < \Delta_d(0)$. This is intuitive; with some VAT evaded on sales, output VAT becomes less of a burden. It then follows that T is increasing in ν_R i.e. voluntary registration is more likely, the greater the opportunities for evasion, as measured by ν_R.

As regards bunching, evasion has two opposing effects. First, evasion relaxes the constraint imposed by the VAT threshold, as the tax authority only observes $1 - \nu_N$ of B2C sales, and so the firm can in fact produce over the threshold without registering. Second, as just discussed, if $\nu_R > 0$, evasion makes registration less costly, because output VAT becomes less of a burden.

Both of these effects appear formally as follows. With evasion, we show below that the term T in the bunching equation (18) is replaced by

$$\hat{T}(\nu_N, v_R) = \frac{(1-\nu_N)\lambda + (1-\lambda)A_B}{\lambda + (1-\lambda)A_B} T(\nu_R) \tag{B.16}$$

and in particular, positive bunching will occur when $\hat{T}(\nu_N, \nu_R)$ is less than 1.[29] An increase in ν_N has the expected effect of lowering $\hat{T}$ and thus increasing bunching, as a higher ν_N makes it less costly for the firm to hold its observed sales under the threshold. An increase in ν_R has the opposite effect of reducing bunching, via the fact that $T(\nu_R)$ rises; this captures the effect that evasion reduces the burden of output VAT.

Note that, with evasion, the qualitative effects of λ and ω on T do not change, and so our predictions about the determinants of voluntary registration do not change; this is clear by inspection from (B.15). This is also true of bunching; it is seen by inspection that $\hat{T}$ is decreasing in λ, and increasing in ω, as is T. So, our key empirical predictions are robust to the presence of evasion. We can summarize as follows:

Proposition b.1 *An increase in evasion by non-registered firms, ν_N, increases bunching. An increase in evasion by registered firms, ν_R, raises the likelihood of voluntary registration and reduces bunching. Moreover, evasion does not affect our qualitative predictions about the effects of λ, the fraction of B2C sales, and input-cost ratio, ω, on voluntary registration and bunching.*

[29] For a formal proof, see below.

Derivation of the Bunching Equation with Evasion

The proof follows the proof of Proposition 4 in the paper, with the following changes. First, we define $A_0,\ A_1$ as

$$A_0(\nu) = \lambda(1-\nu_N) + (1-\lambda)A_B,\ A_1(\nu) = \lambda(1+(1-\nu_R)t)^{-e} + (1-\lambda)A_B$$

As in (A.2) in the paper, any firm that registers has maximized profit:

$$\pi(1;a) = A_1(\nu_R)\left(\frac{e}{1-e}\right)^{-e}\frac{a^{e-1}}{1-e} \tag{B.17}$$

Next, the payoff from being on the VAT threshold for an $a-$type when constrained is now

$$\pi(0;a) = (s^*+\nu px) - \frac{1+\Delta_c}{a}\frac{(s^*+\nu px)}{p} \tag{B.18}$$

This is because the firm can actually produce and sell $s^*+\nu px$ with a threshold s^* because sales νpx are "cash" and thus not observable by the tax authority. Solving for p from the definition that non-concealed sales must be equal to s^* i.e. $((1-\nu_N)\lambda+(1-\lambda)A_B)p^{1-e} = s^*$, we get:

$$p = \left(\frac{s^*}{(1-\nu_N)\lambda+(1-\lambda)A_B}\right)^{-1/(e-1)}$$

Combining this with the fact that $x = \lambda p^{1-e}$, we get

$$s^*+\nu px = s^*\frac{\lambda+(1-\lambda)A_B}{(1-\nu_N)\lambda+(1-\lambda)A_B} \equiv \mu s^*$$

Substituting this back into (B.18), and setting $a = a^*+\Delta a^*$, we get

$$\pi(0;a^*+\Delta a^*) = \mu\left(s^* - \frac{1+\Delta_c}{(A_0(v_N))^{1/(e-1)}(a^*+\Delta a^*)}(s^*)^{e/(e-1)}\right) \tag{B.19}$$

Also observed non-cash sales map into type by

$$s^*+\Delta s^* = A_0(v_R)\left(\frac{e\,(1+\Delta_c)}{e-1}\right)^{1-e}(a^*+\Delta a^*)^{e-1} \tag{B.20}$$

Combining (B.19) and (B.20), we get:

$$
\begin{aligned}
\pi(0; a^* + \Delta a^*) &= \mu\left(s^* - \frac{1+\Delta_c}{(A_0(v_N))^{1/(e-1)}(a^*+\Delta a^*)}(s^*)^{e/(e-1)}\right) \qquad \text{(B.21)} \\
&= \mu\left(s^* - (1+\Delta_c)\,(s^*)^{e/(e-1)}\left(\frac{e-1}{e\,(1+\Delta_c)}\right)(s^*+\Delta s^*)^{1/(1-e)}\right) \\
&= \mu\left(s^* - (s^*)^{e/(e-1)}\left(\frac{A_0(v_R)}{A_0(v_N)}\right)^{1/(e-1)}\left(\frac{e-1}{e}\right)(s^*+\Delta s^*)^{1/(1-e)}\right)
\end{aligned}
$$

Now using (B.20) in (B.17), we get:

$$
\begin{aligned}
\pi(1; a^* + \Delta a^*) &= A_1(v_R)\left(\frac{e}{1-e}\right)^{-e}\frac{1}{1-e}(a^*+\Delta a^*)^{e-1} \qquad \text{(B.22)} \\
&= \frac{A_1(v_R)}{A_0(v_R)}\left(\frac{e}{1-e}\right)^{-e}\frac{1}{1-e}\left(\frac{e(1+\Delta_c)}{e-1}\right)^{e-1}(s^*+\Delta s^*) \\
&= \frac{A_1(v_R)(1+\Delta_c)^{e-1}}{A_0(v_R)}\frac{(s^*+\Delta s^*)}{e} \\
&= T(v_R)\frac{(s^*+\Delta s^*)}{e} \qquad \text{(B.23)}
\end{aligned}
$$

So, using (B.22),(B.23), the indifference condition $\pi(1; a^* + \Delta a^*) = \pi_0(0; a^* + \Delta a^*)$ becomes

$$
s^* - (s^*)^{e/(e-1)}\left(\frac{e-1}{e}\right)(s^*+\Delta s^*)^{1/(1-e)} - \frac{T(v_R)}{\mu}\frac{(s^*+\Delta s^*)}{e} = 0 \qquad \text{(B.24)}
$$

After some simplification of (B.24) (divide through by s^*, then $1 + \frac{\Delta s^*}{s^*}$, and multiply by e) we get (18) in the paper with the tax term $\frac{T(v_R)}{\mu}$ as required. □

B.1 Compliance Costs

Suppose that there is a fixed compliance cost $\Gamma > 0$ for the firm if it registers. What will the effect be on the equilibrium? If the voluntary registration condition (16) above does *not* hold, so we are in the bunching regime, then firms continue to bunch, but the upper end of the bunching interval is now described by

$$
\pi(1; a^* + \Delta a^*) - \Gamma = \pi(0; a^* + \Delta a^*) \qquad \text{(B.25)}
$$

Clearly, the larger Γ, the higher is the upper end of the bunching interval in (B.25), as firms now have an incentive not only to avoid charging the output VAT but also the compliance cost. If the voluntary registration condition condition (16) above *does* hold, then there are two possibilities. If Γ is small enough, then even the least productive $\underline{a}$-firm will choose to register voluntarily and so Proposition 16 continues to hold. If Γ is larger, then low-productivity firms start to bunch.

C The Optimal VAT Threshold

In this section, we show how a formula for the optimal VAT threshold can be derived for our model. Following Keen and Mintz (2004), we assume that any firm that is registered has to pay a compliance cost $\Gamma > 0$ and the revenue authority incurs an administration cost $A > 0$. To keep things, simple, we will also assume that $e_C = e_B = e$, so $p_C(a) = p_B(a) = p(a)$.

As a first step, it is helpful to rule out parameter values under which all firms choose to register voluntarily. As discussed in Section 4.3, this situation arises when Γ is small enough, and condition (16) in the paper holds. Then, even the least productive $\underline{a}$-firm will choose to register voluntarily, so the location of the registration threshold s^* becomes irrelevant to firm behavior, government revenue and welfare. Thus, we assume that either; (i) (16) does not hold, or (ii) (16) does hold, and Γ is large enough to cause some firms to bunch.

Let the firm at the top of the bunching interval be $a^* + \Delta a^* \equiv \hat{a}$ for convenience. Then $\hat{a}$ is defined by condition (B.25) in the paper. Note for future reference that $\hat{a}$ is strictly increasing in s^*, as the higher the threshold, the high the productivity of the firm that is just willing to cut output to stay below the threshold. We do not need to solve for the details of this relationship in what follows, we just need to recall that $\frac{\partial \hat{a}}{\partial s^*} > 0$.

We focus on the choice of the VAT threshold to maximize tax revenue, to bring out the basic issues. The VAT base, which is taxed at rate t, is

$$B = PY + \int_{\hat{a}}^{\overline{a}} p(a)x(a)da + \omega \int_{\hat{a}}^{\overline{a}} \frac{x(a) + y(a)}{a} da \tag{C.26}$$

This is composed of three terms; the value of final sales to the household of the large firm, the value of final sales of the registered small firms, and the value of inputs purchased by non-registered small firms. Revenue is therefore

$$R = tB - (\overline{a} - \hat{a})A \tag{C.27}$$

where A is the administration cost per firm of implementing the VAT, and $(\overline{a} - \hat{a})$ is the measure of firms that are registered.

To further study this expression, we note that the price $p(a)$ charged by the firms in equilibrium is as follows:

$$p(a)=\begin{cases}\frac{e}{e-1}\frac{1}{a}, & a>\hat{a}\\ \left(\frac{\Omega}{s^*}\right)^{1/(e-1)}, & a^*\leq a\leq\hat{a}\\ \frac{e}{e-1}\frac{1+\omega t}{a}, & a<a^*\end{cases}\tag{C.28}$$

where $\Omega=\lambda+(1-\lambda)A_B$. The explanation is as follows. For firms $a>\hat{a}$, the price is a fixed mark-up over marginal cost of $1/a$. The same is true for firms $a<a^*$, except that now marginal cost is $(1+\omega t)/a$, due to embedded VAT. For firms in the bunching interval, the price is set to just satisfy demand, while making the total value of sales equal to s^* i.e. $p(a)$ must satisfy $p(a)(x(a)+y(a))=s^*$; using (12), (13), and solving for p gives the formula (C.28). Finally, in what follows, define $p(s^*)\equiv\left(\frac{\Omega}{s^*}\right)^{1/(e-1)}$.

Next, we note that from (6), (8) in the paper, and using $e=\gamma$, gives

$$PY=(1-\lambda)(1+t)^{-e}\left(\frac{e}{e-1}C\right)^{1-e}=(1-\lambda)A_B\int_{\underline{a}}^{\overline{a}}(p(a))^{1-e}\,da\tag{C.29}$$

So, combining (C.28), (C.29):

$$PY=(1-\lambda)A_B\left[(p(s^*))^{1-e}(\hat{a}-a^*)+\int_{\hat{a}}^{\overline{a}}(p(a))^{1-e}\,da+\int_{\underline{a}}^{a^*}(p(a))^{1-e}\,da\right]\tag{C.30}$$

Also, using the fact that for $a^*\leq a\leq\hat{a}$, $x(a)+y(a)=s^*/p(s^*)$, we can write:

$$\int_{\underline{a}}^{\hat{a}}\frac{x(a)+y(a)}{a}da=\frac{s^*}{p(s^*)}\int_{a^*}^{\hat{a}}\frac{1}{a}da+\int_{\underline{a}}^{a^*}\frac{x(a)+y(a)}{a}da\tag{C.31}$$

So, combining (C.30),(C.31) and (C.26), we get:

$$\begin{aligned}B\ &=\ (1-\lambda)A_B\left[(p(s^*))^{1-e}(\hat{a}-a^*)+\textstyle\int_{\hat{a}}^{\overline{a}}(p(a))^{1-e}\,da+\int_{\underline{a}}^{a^*}(p(a))^{1-e}\,da\right]\\ &+\ \textstyle\int_{\hat{a}}^{\overline{a}}p(a)(x(a))da+\omega\frac{s^*}{p(s^*)}\int_{a^*}^{\hat{a}}\frac{1}{a}da+\omega\int_{\underline{a}}^{a^*}\frac{x(a)+y(a)}{a}da\end{aligned}\tag{C.32}$$

Now, consider the effect of a change in the threshold on the tax base B. When s^* changes, it changes the endpoints of the bunching interval $a^*,\hat{a}$. It also changes the prices via (C.28). Some of these effects will be zero, because the price function $p(a)$ is continuous in a, and also B is continuous in $a^*,\hat{a}$ with one exception; at $\hat{a}$, the terms in the integrands in the second and third terms of (C.32) are generally not equal. In fact, the first is the value of final sales for the $\hat{a}$ firm, and the second is the overall cost of production for the $\hat{a}$ firm, which will be smaller. So, differentiating (C.32), and integrating the term $\int_{a^*}^{\hat{a}} \frac{1}{a} da$ explicitly, we get

$$\begin{aligned} \frac{\partial B}{\partial s^*} &= \left(\frac{\omega}{\hat{a}}\frac{s^*}{p(s^*)} - p(\hat{a})x(\hat{a})\right)\frac{\partial \hat{a}}{\partial s^*} \\ &+ p'(s^*)\left[(1-\lambda)A_B(1-e)(p(s^*))^{-e}(\hat{a}-a^*) - \omega\frac{s^*}{(p(s^*))^2} ln\left(\frac{\hat{a}}{a^*}\right)\right] + \frac{\omega}{p(s^*)} ln\left(\frac{\hat{a}}{a^*}\right) \end{aligned} \tag{C.33}$$

Now from (C.28) and (12), it is easy to establish the following:

$$p'(s^*) = \frac{p(s^*)}{s^*(1-e)},\ p(\hat{a})x(\hat{a}) = \lambda p(\hat{a})^{1-e} = \frac{\lambda s^*}{\Omega} \tag{C.34}$$

Combining (C.33), (C.34), after some simplification, we get

$$\begin{aligned} \frac{\partial B}{\partial s^*} &= \left(\frac{\omega}{\hat{a}}\frac{s^*}{p(s^*)} - (1-\beta)s^*\right)\frac{\partial \hat{a}}{\partial s^*} + \beta(\hat{a}-a^*) \\ &+ \frac{e}{(e-1)}\frac{\omega}{p(s^*)} ln\left(\frac{\hat{a}}{a^*}\right) \end{aligned} \tag{C.35}$$

where $\beta \equiv \frac{(1-\lambda)A_B}{\lambda+(1-\lambda)A_B}$ is the share of demand for the bunching firms that is B2B.

Now using (C.27), we see that the optimal threshold for maximizing revenue is characterized by

$$\frac{\partial R}{\partial s^*} = t\frac{\partial B}{\partial s^*} + A\frac{\partial \hat{a}}{\partial s^*} = 0 \tag{C.36}$$

Combining (C.35) and (C.36), we get the FOC that defines the optimal threshold:

$$\frac{\partial R}{\partial s^*} = \left(t\frac{\omega}{\hat{a}}\frac{s^*}{p(s^*)} - t(1-\beta)s^* + A\right)\frac{\partial \hat{a}}{\partial s^*} + t\beta(\hat{a}-a^*) + t\frac{e}{(e-1)}\frac{\omega}{p(s^*)} ln\left(\frac{\hat{a}}{a^*}\right) = 0 \tag{C.37}$$

To interpret this, and make the link to Keen and Mintz (2004), we proceed as follows. In deriving their simple analytical formula, equation (2) in their paper, Keen and Mintz assume no behavioral responses i.e. no bunching (which can be captured here by $e = 0$), and they also assume no B2B sales i.e. $\beta = 0$. Making these two simplifications in (C.37), we can write

$$\frac{\partial R}{\partial s^*} = (-ts^*v + A)\frac{\partial \hat{a}}{\partial s^*} = 0,\ v = 1 - \frac{\omega}{\hat{a}p(s^*)} \tag{C.38}$$

where v is the share of value added in sales for the $\hat{a}$ firm. Solving (C.38) for s^* gives

$$s^* = \frac{A}{tv} \tag{C.39}$$

This is *exactly* the formula in Keen and Mintz (2004) for the threshold in the special case where the government maximizes revenue[30] This formula is very intuitive; the higher are administrative costs A, the higher the threshold should be, and the higher is the tax rate t or the share of value-added (the tax base) for the firm at the threshold, the lower the threshold will be, as reducing it will increase revenue proportionately to tv.

More generally, we can solve (C.37) for s^* to get a generalization of the Keen and Mintz formula:

$$s^* = \frac{A + t\left(\beta(\hat{a} - a^*) + \frac{e}{(e-1)}\frac{\omega}{p(s^*)} ln\left(\frac{\hat{a}}{a^*}\right)\right) / \frac{\partial \hat{a}}{\partial s^*}}{t\,(v - \beta)} \tag{C.40}$$

This formula tells us that introducing either B2B sales tends to raise the threshold. [31] This is straightforward to see mathematically from (C.40): if not all sales are B2C i.e. $\beta > 0$, then first, a positive term in β appears in the numerator of (C.40), and second, a negative term in β appears in the denominator. Both of these make s^* higher.

The intuition is as follows. First, an increase in s^* increases PY by increasing the prices that the small firms who are bunching charge for inputs to the large firm, and this raises B. This is captured by the term in β in the numerator of (C.40), which is proportional to the mass of bunching small firms, $\beta(\hat{a} - a^*)$.

Second, an increase in s^* raises a^* and therefore the number of firms who bunch. This in turn decreases the net VAT paid by bunching firms, as more firms now below the threshold. But, the size of this effect depends on B2B sales. Specifically, this reduction is proportional to the value of B2C sales minus the value of inputs used to produce *all* sales. So, the higher are B2B sales, the smaller this effect is, and thus the lower the revenue cost of raising s^*. This explains the term in β in the denominator in (C.40).

D Evidence of Bunching via Turnover Misreporting

In this section, we provide some suggestive evidence on the extent of bunching due to turnover misreporting. When bunching is due to a decrease in real output, we expect companies to

[30]This is their equation (2), in the limit as δ, the weight on revenue, goes to infinity, which is $s^* = A/tv$ in our notation.

[31]It can also be seen that the effect of a behavioral response i.e. $e > 0$ on the threshold is positive. This is most easily seen by setting $\beta = 0$ in (C.40) to get the special case of only B2C sales. Then, the term in $e/(e-1)$ is positive. This is a somewhat different finding from Keen and Mintz (2004), who find an ambiguous effect of the behavioral response on the threshold

reduce their input costs in proportion, so that the distribution of input-cost ratio for non-registered companies should be smooth around the VAT notch. When bunching is due to turnover misreporting, we conjecture that the non-registered companies are less likely to under-report their input costs and wage expenses. Both costs are deductible for corporation taxes and the latter is subject to third-party reporting. In other words, the gain from under-reporting the deductible costs is considerably smaller than the gain from under reporting the turnover to avoid VAT registration. If the majority of companies bunch via turnover misreporting, we would expect to see a higher average input-cost ratio for the non-registered group just below the VAT notch, relative to that for the registered group.

Figure A.9 pools all observations in the sample period and plots the distribution of average input-cost ratio for registered and non-registered companies in £1,000 turnover bins, respectively. In Panel A, the input-cost ratio is salary exclusive and represents the share of direct cost of sales relative to total turnover. The solid blue line shows the average input cost relative to sales for registered companies within each turnover bin of £1,000 normalized by the current-year VAT notch, and the dashed blue line shows the average input cost ratio for the unregistered companies. Consistent with the theory, voluntary registers incur a much larger input cost as indicated by their average input-cost ratio which is consistently larger than that for the non-registered companies below the VAT notch. On the other hand, there is no evident increase in the average input-cost ratio just below the VAT notch for the non-registered group. The distribution is relatively smooth and continues to increase with turnover above the VAT notch.

In comparison, Panel B plots the distribution of average input-cost ratio *inclusive* of salary, for registered and non-registered companies, respectively. There is striking difference between the two input-cost ratio series just below the VAT notch. The two series move in parallel directions until the average input-cost ratio for the non-registered companies starts to increase drastically just below the VAT notch. The sharp increase in the salary-inclusive cost ratio can be partly attributed to the fixed nature of salary cost which takes longer to adjust than variable costs of input. On the other hand, the sharp increase is also consistent with the fact that salary is subject to third-party reporting and thus it is more costly/difficult for small businesses to underreport salary expenses. Overall, Panel A and B in Figure A.9 provide suggestive yet not conclusive evidence that part of bunching is due to turnover misreporting.

E Dynamic Regressions of Voluntary Registration

In this section, we investigate the importance of inertia in driving VAT registration by analyzing the dynamic behavior of firms when they cross the registration and deregistration thresholds. First, we compute a transition probability matrix for firms changing their registration states in Table A.4, which shows the probability of being registered or not registered t years after initially being in a given state.[32] For example, the entry in the first cell of the matrix, indicates that of all the firms that where initially registered in year 2004/05, 82.2% remained registered a year later.

Table A.4 shows that there is considerable persistence in registration status. On the other hand, comparing to the registered firms, this persistence does decline substantially over time for non-registered firms. For example, 81.4% of initially non-registered firms remain unregistered after 5 years, whereas 64% of registered firms remain registered after 5 years. The difference in persistence is due to the fact that the majority of firms are growing over our sample, and so will tend to stay above the registration threshold once they cross it.

Next, we investigate to what extent the registration decision is driven by persistence in turnover versus the costs of changing registration status. To answer this question, we augment equation (19) as follows:

$$\begin{aligned} R_{it} &= \gamma_0 + \gamma_1 R_{i,t-1} + \gamma_2 (1 - R_{i,t-1}) IR_{it} + \gamma_3 R_{it-1} ID_{it} \\ &\quad + \gamma_4 B2C^j_{it} + \gamma_5 ICR_{it} + \gamma_6 L^j_{it} + \gamma_7 D_{it} + \rho_t + \phi_i + \nu_{it} \end{aligned} \tag{E.41}$$

where R_{it} is a dummy indicator that takes value 1 if the firm is currently registered and zero otherwise, as defined previously in Section 6.1. In addition,

$$IR_{it} = \begin{cases} 1, & Y_{it} \geq Z_t \\ 0, & Y_{it} < Z_t \end{cases}, \quad ID_{it} = \begin{cases} 1, & Y_{it} \geq Z'_t \\ 0, & Y_{it} < Z'_t \end{cases},$$

where Z_t, Z'_t are the registration and deregistration thresholds at time t, Y_{it} denotes the current-period turnover, so IR_{it} and ID_{it} are dummy indicators recording whether the firm is above the registration and deregistration thresholds respectively at time t. All the other variables are defined as before, and υ_{it} is the error term. We estimate equation (E.41) in a fixed-effect Probit model, and augment the estimation equation with the initial registration

[32] Changes in the transition probability could also driven by attrition, therefore we focus on a balanced sample of firms that we observe throughout the sample period.

status R_{i0} and the mean characteristics of all the time-varying regressors (Wooldridge, 2005).

So, if firm registration decision was entirely backward-looking and ignores its current turnover relative to Z_t, Z_t', we would expect the coefficients γ_2 to γ_5 to be insignificant. However, we expect most firms to comply with the VAT law. Specifically, we expect firms to register when they are initially not registered and their turnover passes above the threshold. Such a firm has a value 1 for the term $(1 - R_{it-1})IR_{it}$ and so we expect to find a positive γ_2.

The VAT legislation also requires firms to stay registered if they are registered in the previous year and their current turnover is above the deregistration threshold. Such a firm will have a value of 1 for $R_{it-1}ID_{it}$. So, we also expect to find that $\gamma_3 > 0$. On the other hand, if the firm remains registered simply due to the cost of deregistration such that whether crossing the deregistration threshold plays no role in the registration behavior, we would expect to find that $\gamma_3 = 0$.

Finally, we already know from our analysis of voluntary registration that the registration decision is significantly affected by the industry B2C ratio $B2C_{it}^j$, the firm input cost ratio ICR_{it}, and the industry-level Lerner index L_{it}^j, so we expect γ_4, γ_5 and γ_6 to be positive.

Table A.5 reports the full results from estimating equation (E.41) using a fixed-effects Probit model, following Wooldridge (2005). For ease of interpretation, Table A.6 reports the relevant average partial effects, which refer to the effect on the mean probability of registration after averaging the unobserved heterogeneity across all firms in the sample. For example, to calculate the APE of a discrete change of ID from 0 to 1, we first compute the average predicted probability of registration at fixed values of $R_{t-1} = 1, ID_t = 0$ and $R_{t-1} = 1, ID_t = 1$, respectively, across all firms in the sample. We then take the difference between the two average probabilities to obtain the average partial effect of ID. We use a similar procedure to compute the average partial effect of a one-standard-deviation increase in $B2C_{it}^j$, ICR_{it}, and L_{it}^j, noting that the one-standard-deviation increase applies to their mean characteristics in addition to the time-varying values. Column (1) shows the mean predicted probability of VAT registration at fixed value of R_{t-1}, ID_t, IR_t, the mean predicted probability across all firms in the sample and that for one-standard-deviation increase in $B2C_{it}^j$, ICR_{it}, and L_{it}^j. Column (2) shows the average partial effects of these variables by taking the difference in the mean predicted probabilities given the change in their value. For example, for firms that are registered in the previous year, falling below the deregistration threshold lowers the probability of being currently registered by 5 percentage points. Alternatively, for firms that are not registered in the previous year, going above the registration threshold increases their probability of registration by 70.4 percentage points. These findings suggest that the registration decision is not entirely driven by the cost of deregistration or inertia.

Finally, the short-run partial effects of the share of B2C sales and the input cost ratio in the dynamic model are considerably smaller than the static estimates in section 6.1. A one standard deviation increase in the B2C ratio and the Lerner index reduces the probability of registration by 0.1 and 0.25 percentage points, respectively, and there is no significant change in the probability of registration for one standard deviation increase in the input cost ratio.

Overall, we see that while there is a considerable amount of persistence in firm behavior, the registration decision is not entirely driven by inertia due to fixed cost of deregistration. Firms respect the legal registration requirement, and at the same time change their registration decisions in a way that is consistent with profit maximization behavior depicted in Section 3. The probability of registration is also affected significantly by the more fundamental determinants of VAT registration. The positive coefficient estimates and partial effects of $R_{it-1}ID_{it}$, the B2C ratio and the input-cost ratio provides supportive evidence that the VAT registration decision is rational and relates to the fundamental determinants of VAT registration as predicted by the theory in Section 4.

F Data construction

Our data combined two administrative datasets – Corporation Tax returns (CT600) and VAT returns, with two additional data sources - Financial Analysis Made Easy (FAME), and annual sector-level statistics on the share of sales to final consumers (B2C ratio) based on the Office for National Statistics (ONS) Input-Output tables. This section describes how we construct our sample.

We first link CT600 corporate tax returns and FAME company accounts using the company identifier (93% of records in CT600 were merged with FAME with 12 months accounting period), resulting in 7,914,902 company-year records covering April 1, 2004-April 4, 2015. This includes companies that have (i) non-missing turnover, (ii) a single CT600 account in one accounting year with full 12-months period, (iii) non-negative fixed assets (measured in FAME), (iv) do not engage mainly in overseas activities based on the HMRC trade classification, or (v) is not part of a company group. Our sample excludes non-profit organizations/companies – that includes charitable organization, industrial and provident society and company limited by guarantee; we also exclude LLPs, public investment trusts, and companies in the financial services industry.

We then merge the CT600-FAME data with the VAT returns, excluding companies in the VAT record that (i) has registered as part of a group; (ii) has the same registration and deregistration date; (iii) has non-corporation status, or (iv)where there is measurement error – when input VAT is greater than the input value, or output VAT tax rate is greater than the statutory rate.

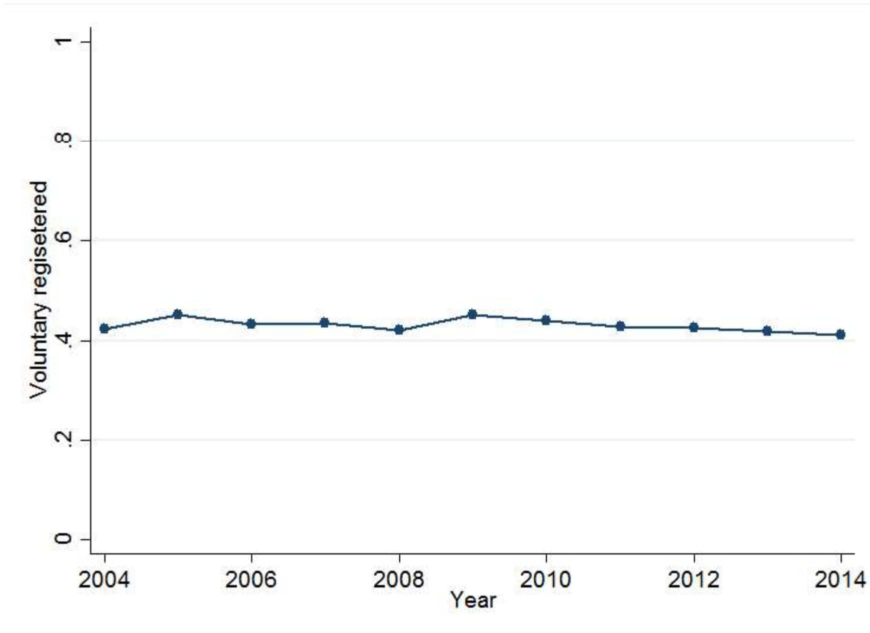
We further link the CT600-FAME-VAT data with annual B2C ratio from ONS at 2-digit SIC (2003) industry level, for industries which the B2C ratios are available from the UK Input-Output tables - that consists of 6,536,170 company-year observations. Our sample for analysis includes observations with yearly turnover between £10,000 to £200,000, and consists of 3,461,247 company-year observations with 968,353 unique companies over the period 2004/2005 to 2014/2015.

G Results excluding firms whose primary product or service is zero-rated or exempted

In this Appendix, we present results on bunching and voluntary registration on our sample excluding firms whose primary product or service is zero-rated or exempted. To identify these sectors, we use the list of exempt and zero-rated products provided by HMRC (https://www.gov.uk/guidance/rates-of-vat-on-different-goods-and-services). We aim to be as certain as possible that we are excluding all sectors where there may be zero-rating or exemption, and so if there is considerable zero-rating and/or exemption within a 2-digit sector, we drop all firms in that sector. Following this rule, we exclude from our sample firms in the following sectors (by 2-digit SIC code, SIC2003 in brackets): agriculture (1), water/sewage (41/90), freight and passenger transport (60/61/62), publishing (22), health (85) and education (80). Table A.7 and figures A.10-A.12 below give the results with this smaller sample and are comparable to Table 2 and Figures 3-5 in the paper. It can be seen that the results are quite similar to those in the main text.

Appendix Figures

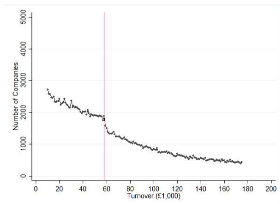
Figure A.1. Registration status of firms below threshold by year



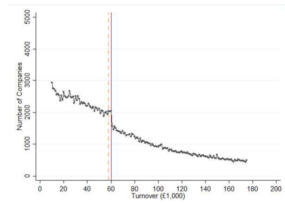
Notes: The figure plots the probability of firms voluntarily registered during 2004/05-2009/10, for firms that are below VAT threshold.

Figure A.2. Annual Turnover Distribution around the Registration threshold, 2004/05-2014/15

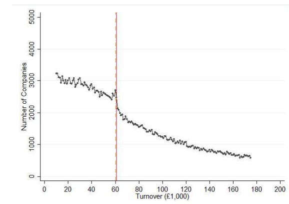
(a) 2004



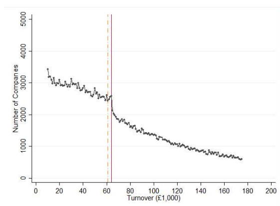
(b) 2005



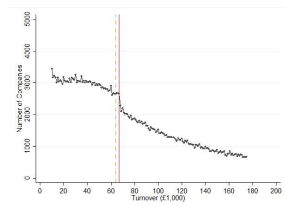
(c) 2006



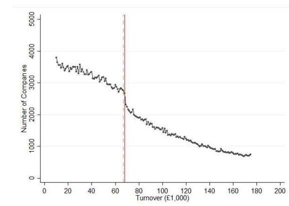
(d) 2007



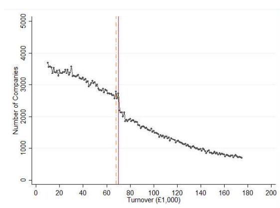
(e) 2008



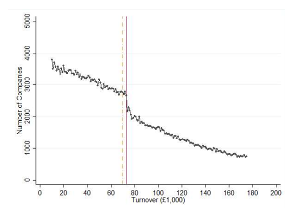
(f) 2009



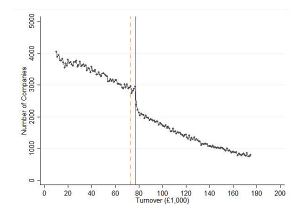
(g) 2010



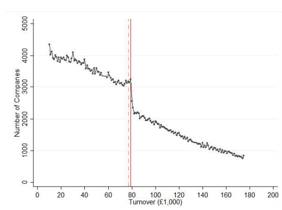
(h) 2011



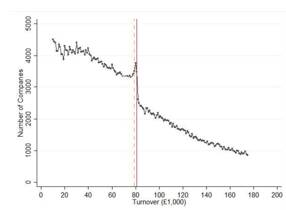
(i) 2012



(j) 2013



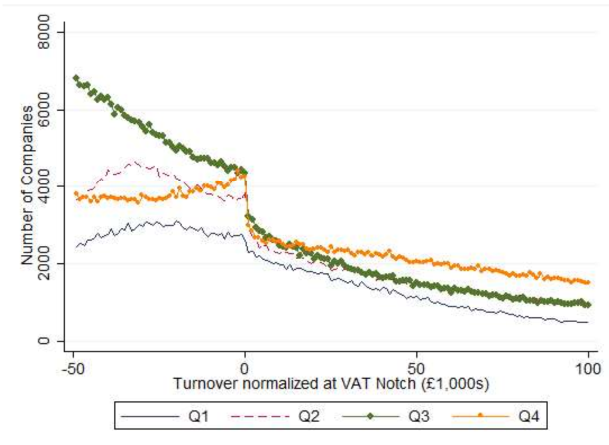
(k) 2014



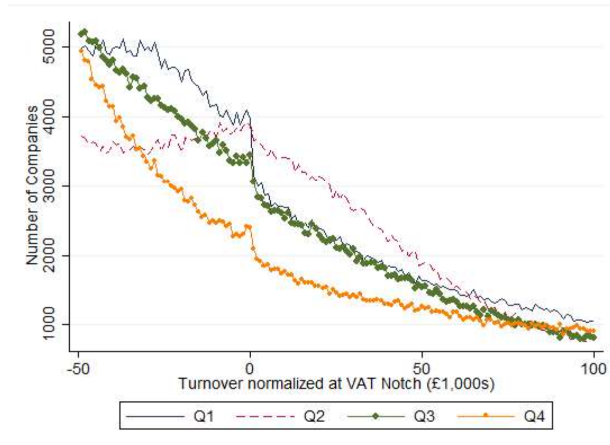
Notes: This figure shows the histogram of companies within the neighborhood of turnover for each year between 2004/05-2014/15. The bin width is £1,000 and the red line denotes the VAT notch. The dash line denotes the VAT notch in the previous year.

Figure A.3. Bunching across Quartiles of the B2C Share Distribution

(a) Binding VAT Threshold



(b) Nonbinding VAT Threshold

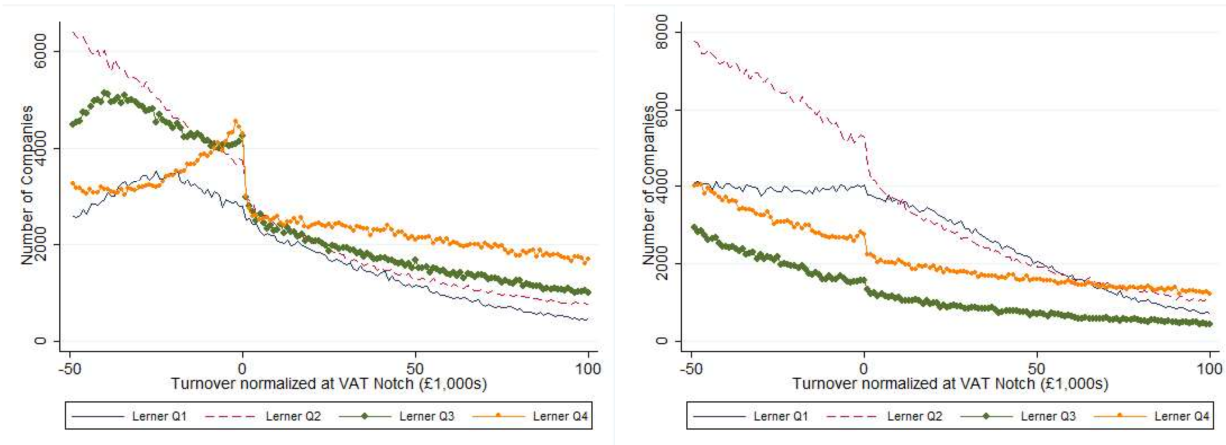


Notes: this figure shows the raw distribution of companies' turnover around the normalized VAT notch across four different quartiles of the distribution of the share of B2C sales for the pooled dataset for period 2004/05-2014/15. Panel (a) shows the distributions for the subset of firms not predicted to register voluntarily. Panel (b) shows the distributions for the subset of firms predicted to register voluntarily, for whom the threshold is nonbinding. The corresponding bunching estimates are reported in Figure 3 in the main text.

Figure A.4. Bunching across Quartiles of the Lerner Index Distribution

(a) Binding VAT Threshold

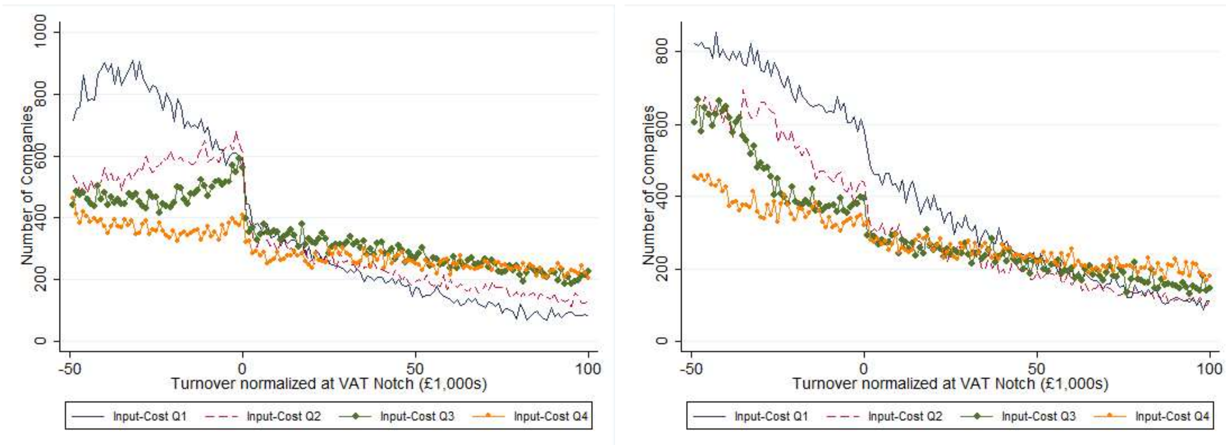
(b) Nonbinding VAT Threshold



Notes: this figure shows the raw distribution of companies' turnover around the normalized VAT notch across four different quartiles of the distribution of the Lerner Index for the pooled dataset for period 2004/05-2014/15. Panel (a) shows the distributions for the subset of firms not predicted to register voluntarily. Panel (b) shows the distributions for the subset of firms predicted to register voluntarily, for whom the threshold is nonbinding. The corresponding bunching estimates are reported in Figure 4 in the main text.

Figure A.5. Bunching across Quartiles of the Input-Cost Ratio Distribution

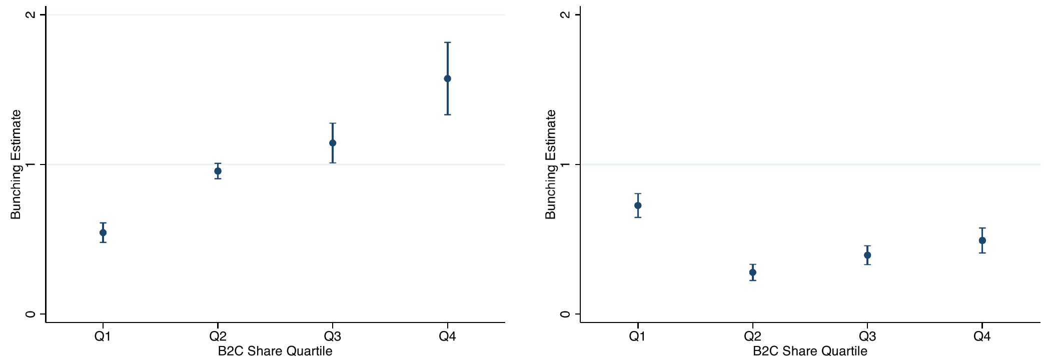
(a) Binding VAT Threshold (b) Nonbinding VAT Threshold



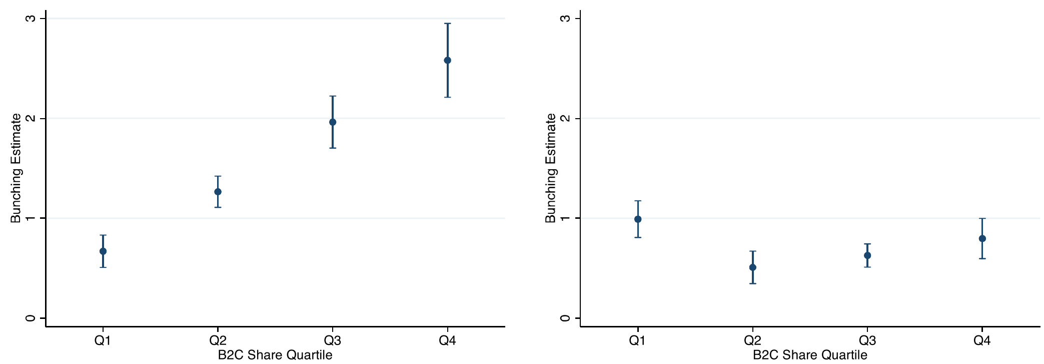
Notes: this figure shows the raw distribution of companies' turnover around the normalized VAT notch across four different quartiles of the distribution of the input-cost ratio for the pooled dataset for period 2004/05-2014/15. Panel (a) shows the distributions for the subset of firms not predicted to register voluntarily. Panel (b) shows the distributions for the subset of firms predicted to register voluntarily, for whom the threshold is nonbinding. The corresponding bunching estimates are reported in Figure 5 in the main text.

Figure A.6. Bunching by Quartiles of B2C, by Lower Bound of Excluded Region

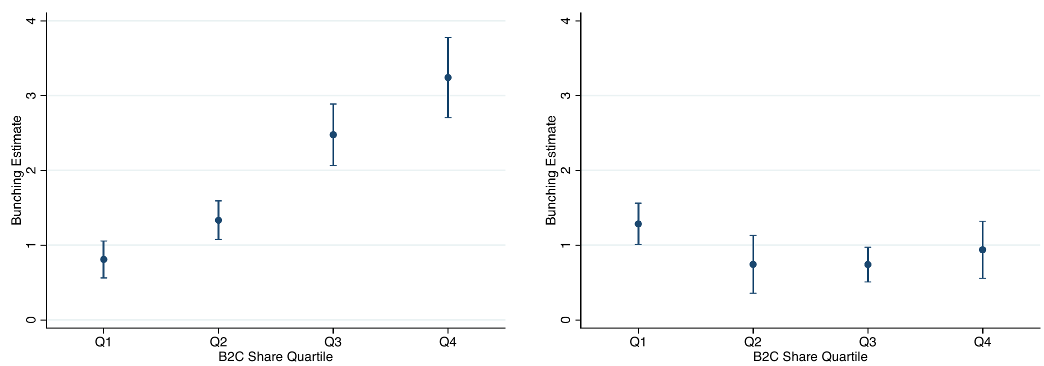
(a) Lower bound = − £6,000



(b) Lower bound = − £12,000



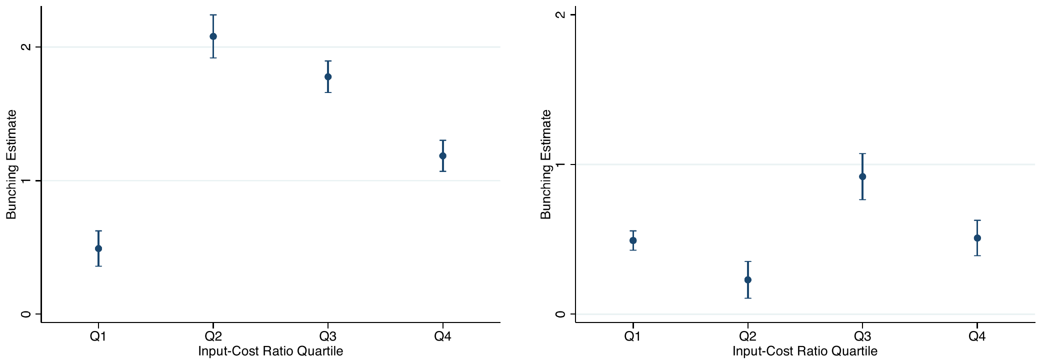
(c) Lower bound = − £18,000



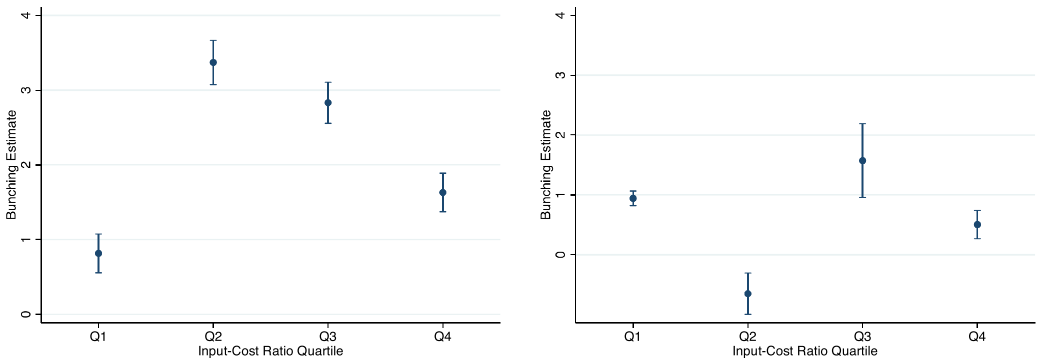
Notes: This figure presents the bunching estimates by quartiles of the distribution of share of B2C sales. The left figure in each panel is constructed for firms predicted not to register voluntarily (i.e., for which the VAT threshold is binding), while the right figure is constructed using only firms predicted to register voluntarily, for which the VAT threshold is non-binding. In panels A (top), B (middle) and C (bottom), we report bunching estimates under the assumption that the excluded range begins £6,000, £12,000 and £18,000 below the VAT threshold, respectively.

Figure A.7. Bunching by Quartiles of Input-Cost Ratio, by Lower Bound of Excluded Region

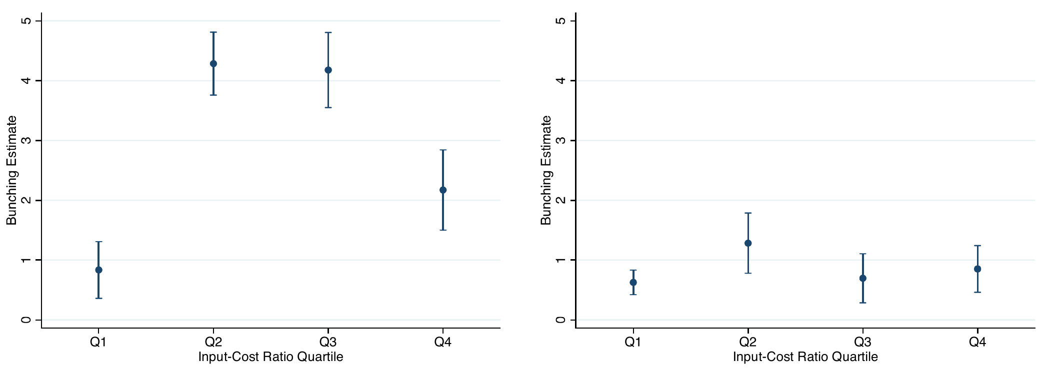
(a) Lower bound = − £6,000



(b) Lower bound = − £12,000



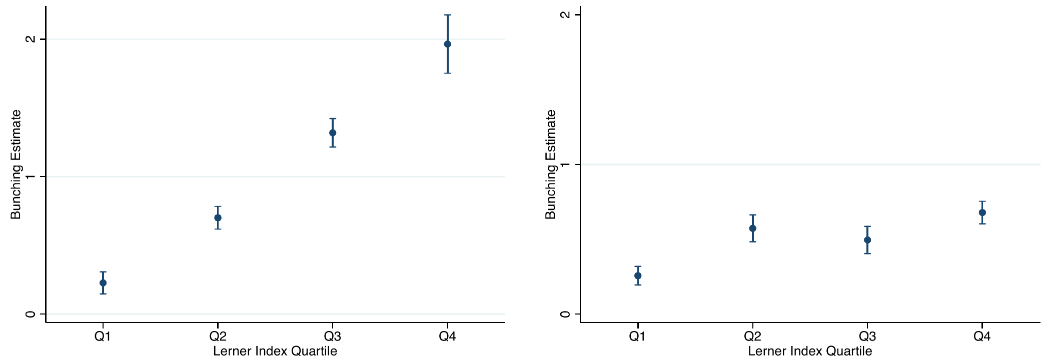
(c) Lower bound = − £18,000



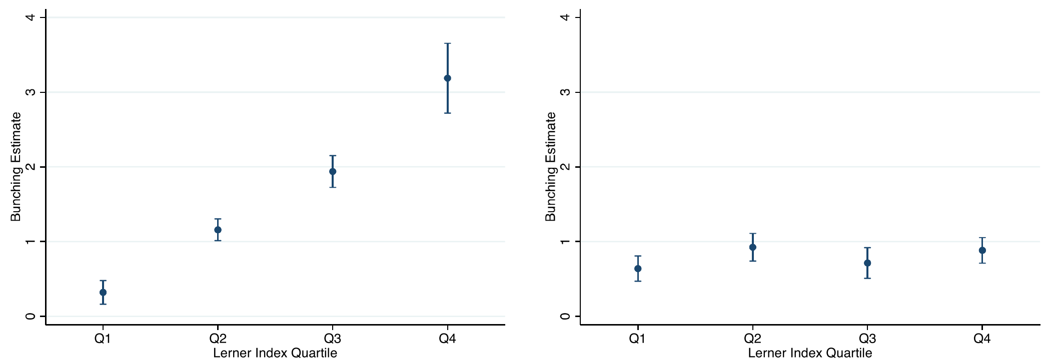
Notes: This figure presents the bunching estimates by quartiles of the distribution of input-cost ratio. The left figure in each panel is constructed for firms predicted not to register voluntarily (i.e., for which the VAT threshold is binding), while the right figure is constructed using only firms predicted to register voluntarily, for which the VAT threshold is non-binding. In panels A (top), B (middle) and C (bottom), we report bunching estimates under the assumption that the excluded range begins £6,000, £12,000 and £18,000 below the VAT threshold, respectively.

Figure A.8. Bunching by Quartiles of Lerner Index, by Lower Bound of Excluded Region

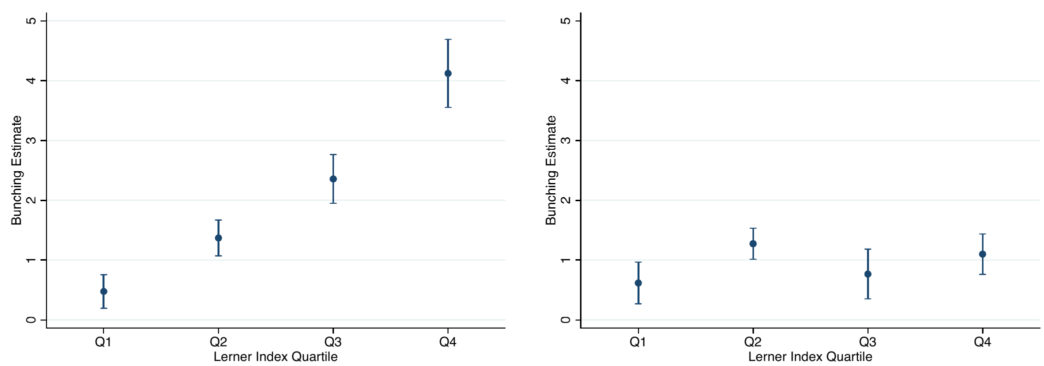
(a) Lower bound = − £6,000



(b) Lower bound = − £12,000



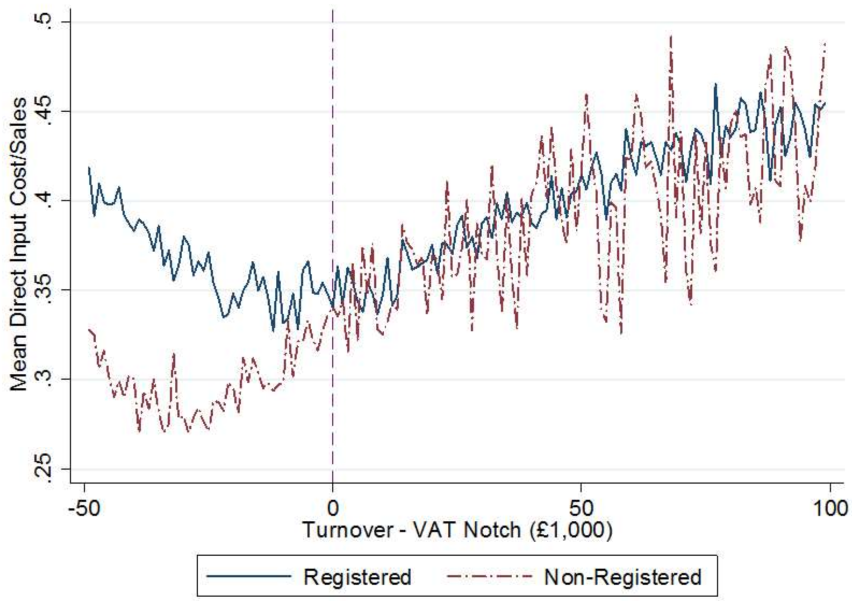
(c) Lower bound = − £18,000



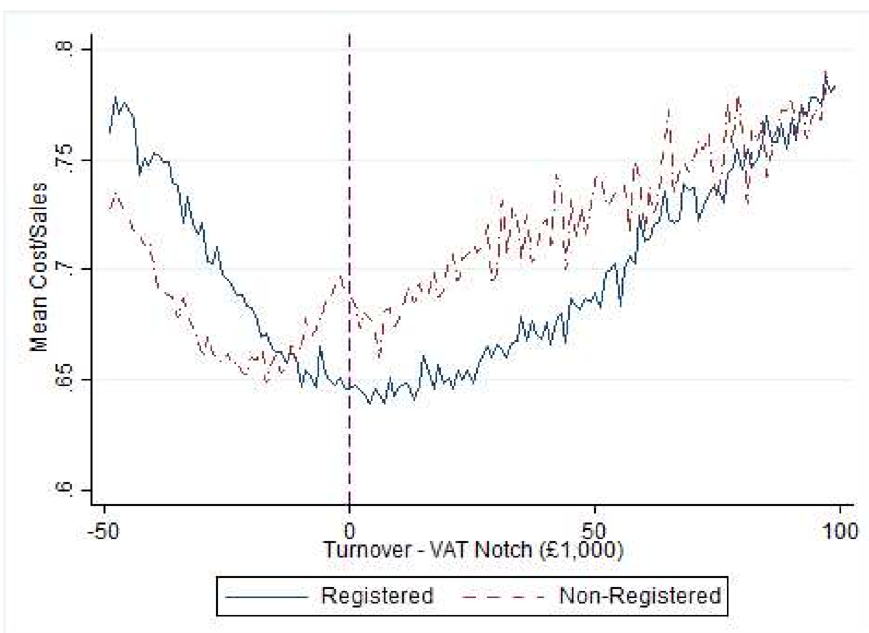
Notes: This figure presents the bunching estimates by quartiles of the distribution of share of B2C sales. The left figure in each panel is constructed for firms predicted not to register voluntarily (i.e., for which the VAT threshold is binding), while the right figure is constructed using only firms predicted to register voluntarily, for which the VAT threshold is non-binding. In panels A (top), B (middle) and C (bottom), we report bunching estimates under the assumption that the excluded range begins £6,000, £12,000 and £18,000 below the VAT threshold, respectively.

Figure A.9. Bunching via Turnover Misreporting

(a) Distribution of Direct Input-Cost Ratio



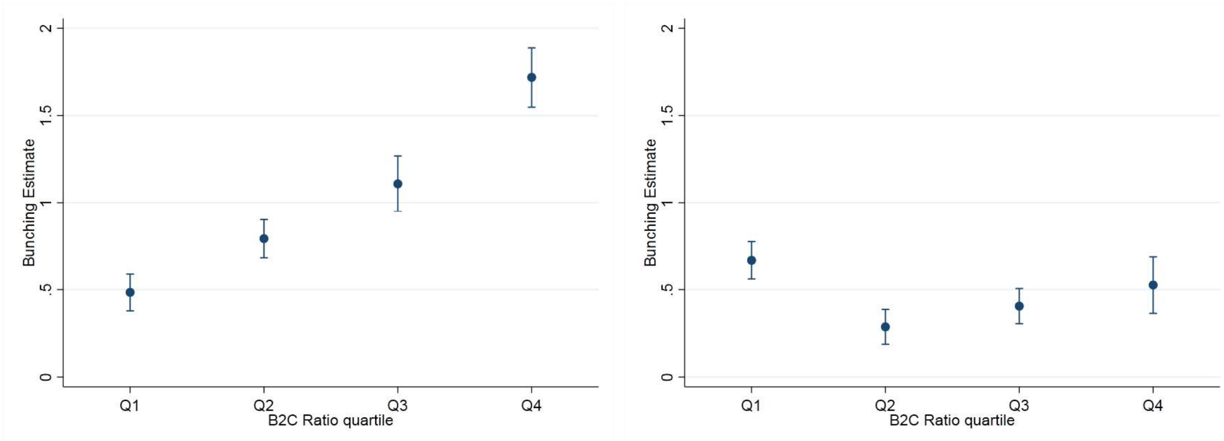
(b) Distribution of Salary-Inclusive Cost Ratio



Notes: the figure plots separately the average input cost ratio for registered and non-registered firms with a turnover in the neighborhood of normalized VAT notch during 2004/05-2009/10. Panel A uses the input cost ratio calculated from FAME and exclude the salary expenses. Panel B uses the input cost ratio calculated from the corporation tax records and includes salary expenses in the overall cost.

Figure A.10. Bunching across Quartiles of the B2C Share Distribution - excluding firms whose primary product or service is zero-rated or exempted

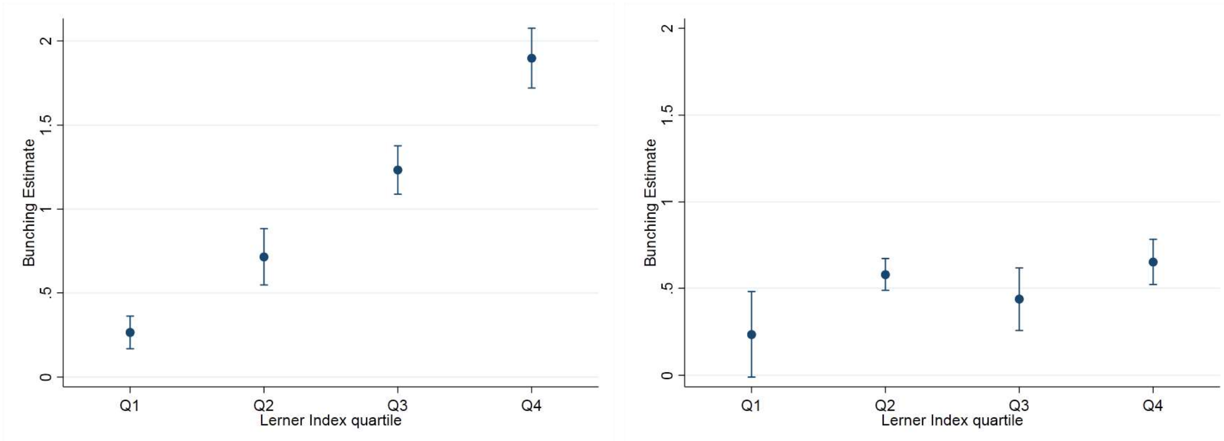
(a) Binding VAT Threshold (b) Nonbinding VAT Threshold



Notes: the figure shows the bunching estimates around the VAT notch across four different quartiles of the distribution of the share of B2C sales. Panel (a) shows the point estimates and 95 percent confidence intervals for the subset of firms not predicted to register voluntarily. Panel (b) shows the estimates for the subset of firms predicted to register voluntarily. Sample exclude firms in the following sectors: agriculture, water/sewage, freight and passenger transport, publishing, health and education.

Figure A.11. Bunching across Quartiles of the Lerner Index Distribution - excluding firms whose primary product or service is zero-rated or exempted

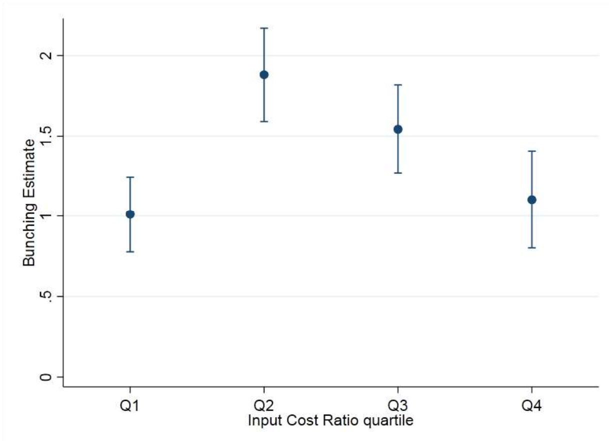
(a) Binding VAT Threshold (b) Nonbinding VAT Threshold



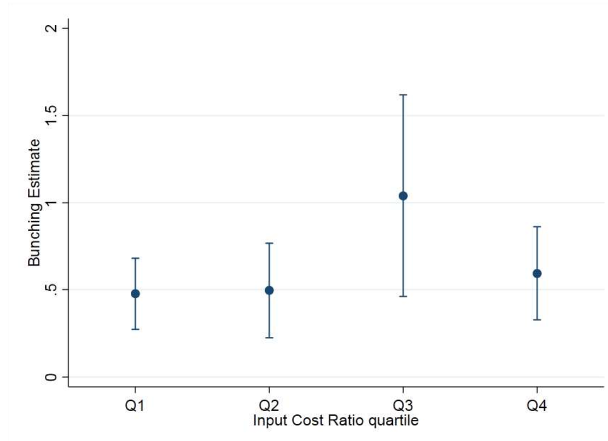
Notes: the figure shows the bunching estimates around the VAT notch across four different quartiles of the distribution of Lerner Index. Panels (a) and (b) differ as in Figure A.10 above. Sample exclude firms in the following sectors: agriculture, water/sewage, freight and passenger transport, publishing, health and education.

Figure A.12. Bunching across Quartiles of the Input-Cost Ratio Distribution - excluding firms whose primary product or service is zero-rated or exempted

(a) Binding VAT Threshold



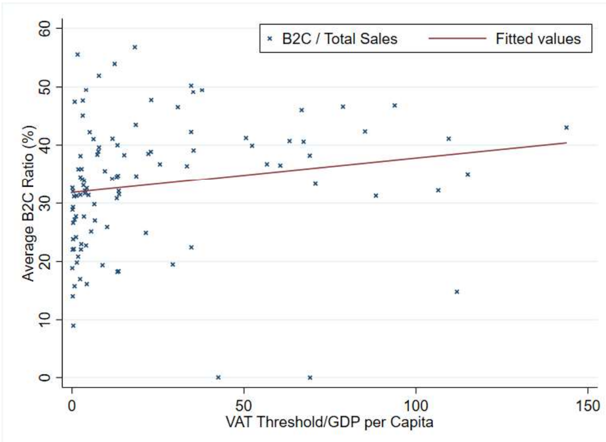
(b) Nonbinding VAT Threshold



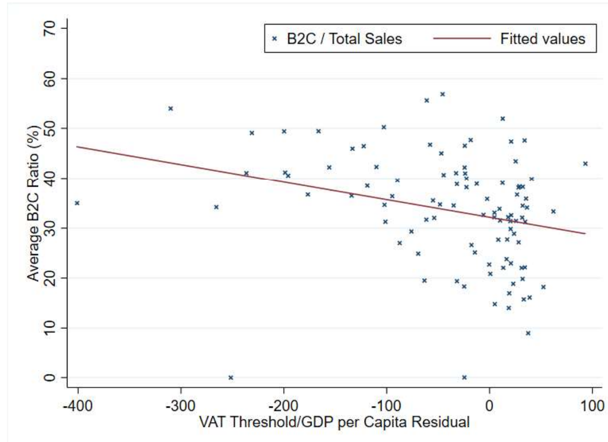
Notes: the figure shows the bunching estimates around the VAT notch across quartiles of the distribution of Input-cost Ratio. Panels (a) and (b) differ as in Figure A.10 above. Sample excludes firms in the following sectors: agriculture, water/sewage, freight and passenger transport, publishing, health and education.

Figure A.13. VAT Threshold and B2C Ratio: Cross-Country Evidence

(a) Raw Correlation



(b) Controlling for the Size of Informal Sector



Notes: the figure shows the correlation between the B2C sales ratio and the VAT threshold across 103 countries in 2017. In panel (a), the VAT threshold is expressed as a fraction of GDP per capita in each country. Panel (b) further controls for the size of the informal sector, by regressing the VAT threshold as a fraction of GDP per capita on the share of agriculture sector in each country and plotting the residual against the B2C sales ratio.

Appendix Tables

Table A.2. Value-Added Tax System in the UK

Fiscal Year	Registration Threshold (£)	Deregistration Threshold (£)	Standard Rate (%)	Flat-Rate Scheme Threshold (£)
2004-05	58,000	56,000	17.5	150,000
2005-06	60,000	58,000	17.5	150,000
2006-07	61,000	59,000	17.5	150,000
2007-08	64,000	62,000	17.5	150,000
Apr 1, 2008-Nov 30, 2008	67,000	65,000	17.5	150,000
Dec 1, 2008-Mar 30, 2009	67,000	65,000	15.0	150,000
Apr 1, 2009-Dec 31, 2009	68,000	66,000	15.0	150,000
Jan 1, 2010-Mar 30, 2010	68,000	66,000	17.5	150,000
Apr 1, 2010-Jan 3, 2011	70,000	68,000	17.5	150,000
Jan 4, 2011-Mar 31, 2011	70,000	68,000	20.0	150,000
2011-2012	73,000	71,000	20.0	150,000
2012-2013	77,000	75,000	20.0	150,000
2013-2014	79,000	77,000	20.0	150,000
2014-2015	81,000	79,000	20.0	150,000

Notes: the table shows changes in the registration threshold, deregistration threshold, Flat-Rate scheme threshold, and VAT standard rate over recent fiscal years. For more information on the UK VAT tax system, see http://www.hmrc.gov.uk/vat/forms-rates/rates/rates-thresholds.htm. For the values of past registration thresholds, see https://www.gov.uk/government/publications/vat-notice-7001-should-i-be-registered-for-vat/vat-notice-7001-supplement–2.

Table A.3. Determinants of Voluntary VAT Registration - other Input cost measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of B2C Sales		-0.192***		-0.004		-0.158***		0.028
		(0.003)		(0.010)		(0.010)		(0.033)
Input-Cost Ratio	0.070***	0.139***	0.067***	0.068***	0.220***	0.287***	0.032***	0.032***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.008)	(0.008)	(0.005)	(0.005)
Lerner Index		-0.341***		-0.226***		-0.696***		-0.163***
		(0.007)		(0.015)		(0.026)		(0.056)
Observations	2405144	2405144	2405144	2405144	156571	156571	156571	156571
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-FE	No	No	Yes	Yes	No	No	Yes	Yes
Cost-Ratio Source	CT600	CT600	CT600	CT600	FAME	FAME	FAME	FAME

Notes: this table presents estimation results from the binary choice model of VAT registration based on equation (19). The dependent variable is the binary indicator of VAT registration status that takes on the value 1 if a firm is voluntarily registered for VAT and zero otherwise. Columns (1)-(2) present results from the linear probability model without firm-fixed effects with input-cost ratio (CT600), and columns (3)-(4) present results by adding firm-fixed effects. Column (5)-(6) present results without firm-fixed effects with input-cost ratio (FAME), and columns (7)-(8) present results by adding firm-fixed effects. Additional firm-level control variables include distance to the registration threshold. *,**,*** denotes significance at 10%, 5% and 1%, respectively. Standard errors are clustered at firm level.

Table A.4. Transition Matrix of VAT Registration Status

	$R_t = 1$ $ID_t = 1$ (1)	$R_t = 1$ $ID_t = 0$ (2)	$R_t = 0$ $IR_t = 1$ (3)	$R_t = 0$ $IR_t = 0$ (4)
$t = 1$				
$R_0 = 1, ID_0 = 1$	82.23%	16.79%	0.30%	0.68%
$R_0 = 1, ID_0 = 0$	19.03%	78.06%	0.16%	2.75%
$R_0 = 0, ID_0 = 1$	9.69%	2.97%	61.57%	25.77%
$R_0 = 0, ID_0 = 0$	2.23%	3.28%	5.60%	88.90%
$t = 2$				
$R_0 = 1, ID_0 = 1$	76.93%	21.13%	0.38%	1.55%
$R_0 = 1, ID_0 = 0$	22.66%	72.31%	0.29%	4.74%
$R_0 = 0, ID_0 = 1$	11.18%	4.97%	53.29%	30.56%
$R_0 = 0, ID_0 = 0$	4.14%	4.99%	5.87%	85.00%
$t = 3$				
$R_0 = 1, ID_0 = 1$	72.67%	24.46%	0.45%	2.42%
$R_0 = 1, ID_0 = 0$	23.53%	69.96%	0.33%	6.19%
$R_0 = 0, ID_0 = 1$	12.29%	5.99%	48.45%	33.26%
$R_0 = 0, ID_0 = 0$	5.24%	6.25%	5.56%	82.94%
$t = 4$				
$R_0 = 1, ID_0 = 1$	68.63%	27.06%	0.63%	3.69%
$R_0 = 1, ID_0 = 0$	23.19%	68.47%	0.40%	7.94%
$R_0 = 0, ID_0 = 1$	12.35%	6.87%	45.13%	35.65%
$R_0 = 0, ID_0 = 0$	6.00%	7.19%	4.88%	81.93%
$t = 5$				
$R_0 = 1, ID_0 = 1$	64.35%	30.68%	0.52%	4.45%
$R_0 = 1, ID_0 = 0$	21.88%	69.23%	0.33%	8.56%
$R_0 = 0, ID_0 = 1$	12.57%	8.01%	41.40%	38.02%
$R_0 = 0, ID_0 = 0$	6.07%	7.96%	4.53%	81.44%

Notes: this table shows in each cell the probability of changing from registration status in year t to year $t + 1$.

Table A.5. Determinants of VAT Voluntary Registration: Probit Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lagged Registration (Rt-1)	3.700***	3.420***	3.973***	3.693***	3.689***	3.681***	3.716***	3.679***	3.702***
	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)	(0.012)	(0.011)	(0.012)
Initial Reg Status (R0)	0.937***	0.869***	0.375***	0.511***	0.756***	0.758***	0.669***	0.756***	0.665***
	(0.023)	(0.022)	(0.016)	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
R*ID		1.838***		1.677***	0.961***	0.971***	0.958***	0.967***	0.957***
		(0.020)		(0.019)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
(1-R)*IR			4.937***	5.200***	4.730***	4.729***	4.790***	4.715***	4.764***
			(0.062)	(0.066)	(0.072)	(0.072)	(0.084)	(0.072)	(0.084)
Distance to Threshold					-0.015***	-0.015***	-0.014***	-0.015***	-0.015***
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Average Distance to Threshold					0.003***	0.003***	0.004***	0.003***	0.003***
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share of B2C Sales						0.094			0.150
						(0.129)			(0.136)
Average Share of B2C Sales						-0.598***			-0.410***
						(0.130)			(0.136)
Input-Cost Ratio							0.229***		0.203***
							(0.020)		(0.020)
Average Input-Cost Ratio							-0.396***		-0.136***
							(0.024)		(0.025)
Lerner Index								0.719***	0.464**
								(0.182)	(0.191)
Average Lerner Index								-2.121***	-1.573***
								(0.187)	(0.197)
Constant	-2.163***	-2.168***	-2.146***	-2.240***	-1.807***	-1.528***	-1.758***	-0.746***	-0.861***
	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)	(0.022)	(0.023)
lnsig2u									
Constant	-0.743***	-0.686***	-2.163***	-1.419***	-0.920***	-0.927***	-1.058***	-0.950***	-1.079***
	(0.030)	(0.029)	(0.077)	(0.044)	(0.032)	(0.032)	(0.036)	(0.033)	(0.036)
Observations	2016377	2016377	2016377	2016377	2016377	2016377	1773902	2016377	1773902
Chi-sq stat B2C						1734.538			373.728
Chi-sq p-val B2C						0			0
Chi-sq stat CostRatio							126.958		18.274
Chi-sq p-val CostRatio							0		0
Chi-sq stat Lerner								2623.417	1201.127
Chi-sq p-val Lerner								0	0
Sigma(u)	0.690	0.710	0.339	0.492	0.631	0.629	0.589	0.622	0.583
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-FE	No	No	No	No	No	No	No	No	No
Cost-Ratio Source							CT600 (new)		CT600 (new)

Notes: this table presents the coefficient estimates from the dynamic estimation of VAT registration in equation (E.41) in a fixed-effect Probit model.

Table A.6. Determinants of VAT Voluntary Registration: Average Partial Effects

Evaluated at:		Mean $\widehat{\Pr(R_t = 1)}$	Average Partial Effect
		(1)	(2)
$R_{t-1} = 1$	$ID_t = 1$	0.987	
$R_{t-1} = 1$	$ID_t = 0$	0.931	0.0557***
			(0.0010)
$R_{t-1} = 0$	$IR_t = 1$	0.989	
$R_{t-1} = 0$	$IR_t = 0$	0.206	0.782***
			(0.0034)
Average in the sample:		0.602	
$B2C + \sigma_{B2C}$		0.600	-0.0028***
			(0.00015)
$ICR + \sigma_{ICR}$		0.603	0.0073***
			(0.00017)
$Lerner + \sigma_{Lerner}$		0.597	-0.0058***
			(0.00017)

Notes: this table presents the partial effects of the key variables of interest from the dynamic estimation of VAT registration in equation (E.41). The partial effects are based on the coefficient estimates reported in column 9 of Table A.5.

Table A.7. Determinants of Voluntary VAT Registration - excluding firms whose primary product or service is zero-rated or exempted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of B2C Sales	-0.182***			-0.048***	-0.029***			0.012
	(0.003)			(0.004)	(0.011)			(0.012)
Input-Cost Ratio		0.131***		0.203***		0.064***		0.065***
		(0.002)		(0.003)		(0.001)		(0.001)
Lerner Index			-0.558***	-0.591***			-0.230***	-0.249***
			(0.006)	(0.008)			(0.016)	(0.017)
Observations	2162231	1931812	2162231	1931812	2162231	1931812	2162231	1931812
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	Yes	Yes	Yes

Notes: this table presents estimation results from the binary choice model of VAT registration based on equation (19). The dependent variable is the binary indicator of VAT registration status that takes on the value 1 if a firm is voluntarily registered for VAT and zero otherwise. Columns (1)-(4) present results from the linear probability model without firm-fixed effects, and columns (5)-(8) present results by adding firm-fixed. The input-cost ratio is the adjusted measure - input-cost ratio (CT600) normalized to match the mean and standard deviation of input-cost ratio (FAME) at industry level. Additional firm-level control variables include distance to the registration threshold. Sample exclude firms in the following sectors: agriculture, water/sewage, freight and passenger transport, publishing, health and education. *,**,*** denotes significance at 10%, 5% and 1%, respectively. Standard errors are clustered at firm level.