Appendix B

This appendix presents the proof of Lemma 1 and the discussion of discontinuities at critical values in the case of indivisible classrooms (case 2).

PROOF OF LEMMA 1

Our strategy for proof is to begin with a preliminary partition of the set of schools and show that within each subset there are two possible optimal integer choices, and that there is a critical value below which the lower integer is optimal and above which the higher integer is optimal. These new critical values become the basis for the new partition described in the statement of the lemma.

1) For a given integer k, let δ_k be the value of λ at which the optimal number of classrooms in Case 1, the divisible-classrooms case, is k. That is, $\delta_k \equiv n^{*-1}(k)$ where $n^*(\lambda)$ is defined by (13c) for $\lambda \geq \alpha$ and by (17c) for $\lambda < \alpha$. As noted above, $n^*(\lambda)$ is continuous and monotonically increasing in λ. This implies that $\delta_{k'} > \delta_k$ if $k' > k$ for any positive integers k, k'. The values δ_k thus form an ordered partition the set of schools.
2) We now show that within the interval $\lambda \in [\delta_k, \delta_{k+1})$, the optimal integer number of classrooms is either k or $k + 1$.
    a) Define functions for profits under the counterfactual assumptions that the class-size cap never binds and that it always binds. That is, let:

    $$\pi^*(n,\lambda) = \pi(p^*(n,\lambda), x^*(n,\lambda), n, \lambda)$$

    where $p^*(n,\lambda)$ and $x^*(n,\lambda)$ are given by (20a) and (20a), the optimal choices when the class-size cap does not bind. Let:

    $$\pi^{**}(n,\lambda) = \pi(p^*(n,\lambda), x^*(n,\lambda), n, \lambda)$$

    where $p^*(n,\lambda)$ and $x^*(n,\lambda)$ are given by (24b) and (24a), the optimal choices when the class-size cap binds. The definition of $\tilde{\pi}(\cdot)$ in (28) can then be restated:

    $$\widetilde{\pi}(n,\lambda) = \begin{cases} \pi^*(n,\lambda) & \text{if} \quad \lambda \leq \beta(n) \\ \pi^{**}(n,\lambda) & \text{if} \quad \lambda > \beta(n) \end{cases} \tag{A15}$$

    b) We now show that $\pi^*(n,\lambda)$, $\pi^{**}(n,\lambda)$ and hence $\widetilde{\pi}(n,\lambda)$ are concave in n for any given λ:
        i) By the envelope theorem, we have:

        $$\frac{\partial \pi^*}{\partial n} = \frac{\partial \mathcal{L}}{\partial n} \tag{A16}$$

        where $\mathcal{L}$ is given by (A1) and $\frac{\partial \pi^*}{\partial n}$ allows p and x to vary (holding λ constant) but $\frac{\partial \mathcal{L}}{\partial n}$ holds p, x and λ constant. Then by (A2b):

        $$\begin{aligned} \frac{\partial^2 \pi^*}{\partial n^2} = \frac{\partial}{\partial n}\left(\frac{\partial \mathcal{L}}{\partial n}\right) &= \frac{\partial}{\partial n}\left(-F_c + \frac{\lambda \Theta^* x^*}{n}\right) \\ &= -\frac{\lambda \Theta^* x^*}{n^2} + \frac{\lambda \Theta^*}{n}\frac{\partial x^*}{\partial n} + \frac{\lambda x^*}{n}\frac{\partial \Theta^*}{\partial n} \end{aligned} \tag{A17}$$

Applying the implicit function theorem (as in (A4)) to (A11a)-(A11c),

$$\text{(A18)} \qquad \begin{pmatrix} \frac{\partial p^*}{\partial n} \\ \frac{\partial x^*}{\partial n} \\ \frac{\partial \Theta^*}{\partial n} \end{pmatrix} = \frac{\lambda}{\mu + \lambda \Psi} \begin{pmatrix} \frac{\lambda \sigma^2_{\theta|\lambda}}{n} \\ \frac{x^* \Psi}{n} \\ \frac{\sigma^2_{\theta|\lambda}}{n} \end{pmatrix}$$

Plugging into (A17),

$$\frac{\partial^2 \pi^*}{\partial n^2} = -\frac{\lambda x \Psi}{n^2 \left(1 + \frac{\lambda \Psi}{\mu}\right)} < 0$$

Hence $\pi^*(n, \lambda)$ is globally concave in n for all λ.

ii) Combining (8), (24a) and (24b), and differentiating twice with respect to n, we have that $\frac{\partial^2 \pi^{**}}{\partial n^2} = -\frac{45\mu}{n} < 0$. Hence $\pi^{**}(n, \lambda)$ is globally concave in n for all λ as well.

iii) Implicitly differentiating (27) with respect to n, we have:

$$\text{(A19)} \qquad \frac{\partial \beta(n)}{\partial n} = \frac{\mu}{n \left(\Psi \ln\left(\frac{T}{45}\right) - \Theta\right)} > 0$$

for all n, where the inequality follows from (22). Hence there is a one-to-one mapping between $\beta(n)$ and n for a given λ. $\beta^{-1}(\lambda)$ is then the value of n at which the class-size cap starts to bind for a given λ. Note that the cap binds for $n < \beta^{-1}(\lambda)$ and does not bind for $n \geq \beta^{-1}(\lambda)$. The definition of $\tilde{\pi}(\cdot)$ in (A15) can then be rewritten:

$$\text{(A20)} \qquad \widetilde{\pi}(n, \lambda) = \begin{cases} \pi^*(n, \lambda) & \text{if} \quad n \geq \beta^{-1}(\lambda) \\ \pi^{**}(n, \lambda) & \text{if} \quad n < \beta^{-1}(\lambda) \end{cases}$$

The function $\pi^*(n, \lambda)$ gives the maximized profit for a given n and λ under one equality constraint, namely $x = d$. The function $\pi^{**}(n, \lambda)$ gives the maximized profit for a given n and λ under two equality constraints, namely $x = d$ and $\frac{x}{n} = 45$. For a given λ, the two points coincide at the point where the optimal class size is 45 even in the absence of the cap—that is, where $n = \beta^{-1}(\lambda)$—and for all other n the function $\pi^{**}(n, \lambda)$ lies under $\pi^*(n, \lambda)$. A standard result in optimization theory is that (if both functions are continuously differentiable) the two curves are tangent at that point (see e.g. Avinash K. Dixit (1976, Ch. 3)). Now consider the curvature of the $\widetilde{\pi}(n, \lambda)$ function. For $n < \beta^{-1}(\lambda)$, $\frac{\partial \widetilde{\pi}}{\partial n}$ is decreasing in n by the concavity of $\pi^{**}(n, \lambda)$. For $n \geq \beta^{-1}(\lambda)$, $\frac{\partial \widetilde{\pi}}{\partial n}$ is decreasing in n by the concavity of $\pi^*(n, \lambda)$. At $n = \beta^{-1}(\lambda)$ the two curves are tangent and $\frac{\partial \pi^*}{\partial n} = \frac{\partial \pi^{**}}{\partial n}$. It follows that $\frac{\partial \widetilde{\pi}}{\partial n}$ is decreasing in n for all n and all λ. Hence $\widetilde{\pi}(n, \lambda)$ is globally concave in n for all λ.

c) Recall the definition of $n^*(\lambda)$, the optimal number of classrooms in the divisible-classroom case, from step 1 of this proof. Since $n^*(\lambda)$ is monotonically increasing in λ, $\lambda \in [\delta_k, \delta_{k+1})$ implies $n^*(\lambda) \in [k, k+1)$. From the concavity of $\widetilde{\pi}(n, \lambda)$ it follows that:

$$\widetilde{\pi}(k,\lambda) > \widetilde{\pi}(k',\lambda)\ \forall\, k' < k \tag{A21}$$
$$\widetilde{\pi}(k+1,\lambda) > \widetilde{\pi}(k',\lambda)\ \forall\, k' > k+1 \tag{A22}$$

That is, within the interval $\lambda \in [\delta_k, \delta_{k+1})$ either k or $k+1$ must be the optimal integer number of classrooms.

3) We now show that in the interval $\lambda \in [\delta_k, \delta_{k+1})$ there is a critical value of λ to the left of which k is the optimal number of classrooms and to the right of which $k+1$ is the optimal number. For $\lambda \in [\delta_k, \delta_{k+1})$, define:

$$\widetilde{\Pi}(\lambda) \equiv \widetilde{\pi}(k+1,\lambda) - \widetilde{\pi}(k,\lambda)$$

Since k is the unique optimal choice of number of classrooms at δ_k in the divisible-classrooms case,

$$\widetilde{\Pi}(\delta_k) = \widetilde{\pi}(k+1,\delta_k) - \widetilde{\pi}(k,\delta_k) < 0$$

Similarly, since $k+1$ is the unique optimal number of classrooms at δ_{k+1} in the divisible-classrooms case,

$$\widetilde{\Pi}(\delta_{k+1}) = \widetilde{\pi}(k+1,\delta_{k+1}) - \widetilde{\pi}(k,\delta_{k+1}) > 0$$

From (A19), we know that $\beta(k) < \beta(k+1)$; the class-size cap starts to bind at a higher value of λ for $n = k+1$ than for $n = k$. Using (A15), the definition of $\widetilde{\Pi}(\lambda)$ can be restated:

$$\widetilde{\Pi}(\lambda) = \begin{cases} \pi^*(k+1,\lambda) - \pi^*(k,\lambda) & \text{if } \lambda \le \beta(k) \\ \pi^*(k+1,\lambda) - \pi^{**}(k,\lambda) & \text{if } \beta(k) < \lambda \le \beta(k+1) \\ \pi^{**}(k+1,\lambda) - \pi^{**}(k,\lambda) & \text{if } \beta(k+1) < \lambda \end{cases}$$

Note that $\pi^*(k,\beta(k)) = \pi^{**}(k,\beta(k))$ and $\pi^*(k+1,\beta(k+1)) = \pi^{**}(k+1,\beta(k+1))$. Hence $\widetilde{\Pi}(\lambda)$ is continuous. Now consider the slope of $\widetilde{\Pi}(\lambda)$ against λ in each interval:

a) $\lambda \le \beta(k)$. The class-size cap binds neither for $n = k$ nor for $n = k+1$.

$$\frac{\partial^2 \pi^*}{\partial n \partial \lambda} = \frac{\partial}{\partial \lambda}\left(\frac{\partial \pi^*}{\partial n}\right) = \frac{\partial}{\partial \lambda}\left(\frac{\partial \mathcal{L}}{\partial n}\right) = \frac{\partial}{\partial \lambda}\left(-F_c + \frac{\lambda x^* \Theta^*}{n}\right) > 0$$

where the second equality follows from (A16), the third equality follows from (A2b), and the inequality follows from (21b) and (23). Hence:

$$\frac{\mathrm{d}\widetilde{\Pi}}{\mathrm{d}\lambda} = \frac{\partial \pi^*(k+1,\lambda)}{\partial \lambda} - \frac{\partial \pi^*(k,\lambda)}{\partial \lambda} > 0$$

b) $\beta(k) < \lambda \le \beta(k+1)$. The class-size cap binds for $n = k$ but not for $n = k+1$. Note that

$$\frac{\partial \pi^*}{\partial \lambda} = \frac{\partial \mathcal{L}}{\partial \lambda} = x^* \Theta^* \ln\left(\frac{nT}{x^*}\right) \tag{A23}$$

where $\mathcal{L}$ is given by (A9), and the first equality follows by the envelope theorem. Similarly,

$$\frac{\partial \pi^{**}}{\partial \lambda} = \frac{\partial \mathcal{L}}{\partial \lambda} = 45n\Theta^* \ln\left(\frac{T}{45}\right) \tag{A24}$$

At $n = k$, the optimal enrollment if there were no class-size cap would be greater than or equal to $45n$; otherwise the cap would not be binding. Hence, comparing (A23) and (A24),

$$\frac{\partial \pi^*(k,\lambda)}{\partial \lambda} \geq \frac{\partial \pi^{**}(k,\lambda)}{\partial \lambda}$$

Using step 3a above,

$$\frac{\partial \pi^*(k+1,\lambda)}{\partial \lambda} > \frac{\partial \pi^*(k,\lambda)}{\partial \lambda} \geq \frac{\partial \pi^{**}(k,\lambda)}{\partial \lambda}$$

Hence:

$$\frac{\mathrm{d}\widetilde{\Pi}}{\mathrm{d}\lambda} = \frac{\partial \pi^*(k+1,\lambda)}{\partial \lambda} - \frac{\partial \pi^{**}(k,\lambda)}{\partial \lambda} > 0$$

c) $\beta(k+1) < \lambda$. The class-size cap binds for both $n = k$ and $n = k+1$. Partially differentiating (A24) and using (A18),

$$\frac{\partial}{\partial n}\left(\frac{\partial \pi^{**}}{\partial \lambda}\right) = 45 \ln\left(\frac{T}{45}\right)\left[\Theta + n\frac{\partial \Theta}{\partial n}\right] > 0$$

Hence

$$\frac{\mathrm{d}\widetilde{\Pi}}{\mathrm{d}\lambda} = \frac{\partial \pi^{**}(k+1,\lambda)}{\partial \lambda} - \frac{\partial \pi^{**}(k,\lambda)}{\partial \lambda} > 0$$

Thus $\widetilde{\Pi}(\lambda)$ is continuous and monotonically increasing in λ for all $\lambda \in [\delta_k, \delta_{k+1})$. Together with the fact that it is negative at δ_k and positive at δ_{k+1}, this implies that there is exactly one critical value, call it ν_k, at which $\widetilde{\Pi}(\nu_k) = 0$. For $\lambda \in [\delta_k, \nu_k)$, k is the optimal integer number of classrooms; for $\lambda \in [\nu_k, \delta_{k+1})$, $k+1$ is optimal.

4) It remains to consider the regions at the extremes of the support of λ. Without loss of generality, let $\underline{j}$ be the largest integer such that $\delta_{\underline{j}} \leq \underline{\lambda}$,[48] and let $\overline{j}$ be the smallest integer such that $\overline{\lambda} \leq \delta_{\overline{j}}$. Within each interval, $[\delta_{\underline{j}}, \delta_{\underline{j}+1})$, $[\delta_{\underline{j}+1}, \delta_{\underline{j}+2})$, ..., $[\delta_{\overline{j}-1}, \delta_{\overline{j}})$, the results from steps 1-3 above hold. Truncate the interval $(\delta_{\underline{j}}, \delta_{\overline{j}})$ at $\underline{\lambda}$ below and $\overline{\lambda}$ above. If $\nu_{\underline{j}} \leq \underline{\lambda}$, then let $\underline{k} = \underline{j}+1$; else if $\underline{\lambda} < \nu_{\underline{j}}$ then let $\underline{k} = \underline{j}$. If $\nu_{\overline{j}-1} < \overline{\lambda}$, then let $\overline{k} = \overline{j}$; else if $\overline{\lambda} \leq \nu_{\overline{j}-1}$, then let $\overline{k} = \overline{j}-1$. Let $\nu_{\underline{k}-1} = \underline{\lambda}$ and $\nu_{\overline{k}} = \overline{\lambda}$. Then

$$\nu_{\underline{k}-1} < \nu_{\underline{k}} < \ldots < \nu_{\overline{k}}$$

form a partition of the set of voucher schools, with the optimal integer number of classrooms equal to $\underline{k}, \underline{k}+1, \ldots, \overline{k}$ between consecutive values, and the lemma is proved.

[48] If $\underline{\lambda} < \delta_1$ then let $\underline{j} = 0$ and $\delta_0 = 0$.

Discontinuities at Critical Values

Consider a given ν_k from Lemma 1, where $\underline{k} < k < \overline{k}$. The fact that $\beta(k+1) > \beta(k)$ (from (A19)) implies that it will never be the case that the class-size cap is non-binding to the left of the critical value but binding to the right of it. There are then three cases to consider:

1) The class-size cap is binding neither to the left nor to the right of the critical value: $\nu_k \leq \beta(k) < \beta(k+1)$. In this case, $\lim_{\lambda\to\nu_k^-} p^*$, $\lim_{\lambda\to\nu_k^-} x^*$ and $\lim_{\lambda\to\nu_k^-} \Theta^*$ (the limits as λ approaches ν_k from the left) are given by (20a), (20b), and (5) with $n = k$. $\lim_{\lambda\to\nu_k^+} p^*$, $\lim_{\lambda\to\nu_k^+} x^*$ and $\lim_{\lambda\to\nu_k^+} \Theta^*$ are given by the same expressions with $n = k+1$. The differences in the left and right limits then have the same signs as the partial derivatives of the variables with respect to n. By (A18), we have immediately that $\frac{\partial p^*}{\partial n} > 0$, $\frac{\partial x^*}{\partial n} > 0$, and $\frac{\partial \Theta^*}{\partial n} > 0$. Hence:

(A25a) $$\lim_{\lambda\to\nu_k^-} p^* < \lim_{\lambda\to\nu_k^+} p^*$$

(A25b) $$\lim_{\lambda\to\nu_k^-} x^* < \lim_{\lambda\to\nu_k^+} x^*$$

(A25c) $$\lim_{\lambda\to\nu_k^-} \Theta^* < \lim_{\lambda\to\nu_k^+} \Theta^*$$

Moreover,

$$\begin{aligned}\frac{\partial}{\partial n}\left(\frac{x^*}{n}\right) &= \frac{1}{n}\frac{\partial x^*}{\partial n} - \frac{x^*}{n^2}\\ &= \frac{x^*}{n^2}\left(\frac{\lambda\Psi}{\mu+\lambda\Psi} - 1\right) < 0\end{aligned}$$

and hence:

(A26) $$\lim_{\lambda\to\nu_k^-} \frac{x^*}{n} > \lim_{\lambda\to\nu_k^+} \frac{x^*}{n}$$

2) The class-size cap is binding to the left of the critical value but not to the right: $\beta(k) < \nu_k < \beta(k+1)$. It is immediate that,

(A27) $$\lim_{\lambda\to\nu_k^-} \frac{x^*}{n} = 45 > \lim_{\lambda\to\nu_k^+} \frac{x^*}{n}$$

Consider x^* and Θ^* in turn:

a) If there were no class-size cap, then for $\lambda \in (\beta(k), \nu_k)$, we would have $\frac{\partial x^*}{\partial \lambda} > 0$. In the presence of the class-size cap, for $\lambda \in (\beta(k), \nu_k)$ we have $\frac{\partial x^*}{\partial \lambda} = 0$. Hence

$$\lim_{\lambda\to\nu_k^-} x^*\bigg|_{cap} < \lim_{\lambda\to\nu_k^-} x^*\bigg|_{no\,cap}$$

If there were no class-size cap, then by (A25b) we would have:

$$\lim_{\lambda \to \nu_k^-} x^* \Big|_{no\ cap} < \lim_{\lambda \to \nu_k^+} x^*$$

Hence:

$$\lim_{\lambda \to \nu_k^-} x^* \Big|_{cap} < \lim_{\lambda \to \nu_k^+} x^*$$

b) Let $z^* = \frac{x^*}{n}$. Recalling the assumption that schools cannot price discriminate, the price term can be brought outside the integral in (5) and Θ^* can then be written:

$$\Theta^* = \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\Omega(\theta)} \left(\frac{T}{z^*}\right)^{\frac{\theta\lambda}{\mu}} g(\theta)\mathrm{d}\theta}{\int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{\Omega(\theta)} \left(\frac{T}{z^*}\right)^{\frac{\theta\lambda}{\mu}} g(\theta)\mathrm{d}\theta} \tag{A28}$$

Note that Θ^* depends only on class size and λ, not on price, or enrollment or the number of classrooms separately. Differentiating,

$$\frac{\partial \Theta^*}{\partial z} = -\frac{\lambda}{\mu z}\sigma^2_{\theta|\lambda} < 0 \tag{A29}$$

Hence (A27) implies:

$$\lim_{\lambda \to \nu_k^-} \Theta^* < \lim_{\lambda \to \nu_k^+} \Theta^* \tag{A30}$$

3) If the class-size cap is binding both to the left and to the right of the critical value: $\beta(k) < \beta(k+1) \leq \nu_k$. It is immediate that:

$$\lim_{\lambda \to \nu_k^-} z^* = \lim_{\lambda \to \nu_k^+} z^* = 45 \tag{A31}$$

and:

$$\lim_{\lambda \to \nu_k^-} x^* = 45k < 45(k+1) = \lim_{\lambda \to \nu_k^+} x^*$$

As (A28) indicates, Θ^* depends only on class size and λ. Hence (A31) implies:

$$\lim_{\lambda \to \nu_k^-} \Theta^* = \lim_{\lambda \to \nu_k^+} \Theta^*$$

Since $\frac{\partial \Theta^*}{\partial \lambda} > 0$ both to the left and to the right of ν_k (refer to (26c)), we have that Θ^* is strictly increasing in λ at ν_k.

In this case, p^* is given by (24a) both to the left and to the right of the critical value. Hence the jump in number of classrooms from k to $k+1$ implies:

$$\lim_{\lambda \to \nu_k^-} p^* > \lim_{\lambda \to \nu_k^+} p^*$$

To summarize, at ν_k, class size is either decreasing or constant in λ, and enrollment and average willingness to pay are always increasing in λ. We have no unambiguous result for how price changes with λ at ν_k.