Near-Optimality of Uniform Co-payments for Subsidies and Taxes Allocation Problems

Retsef Levi
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142, retsef@mit.edu
Georgia Perakis
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142, georgiap@mit.edu
Gonzalo Romero
Rotman School of Management, University of Toronto, Toronto, ON M5S 3E6, gonzalo.romero@rotman.utoronto.ca

We study a subsidies and taxes allocation problem with endogenous market response subject to a budget constraint. The central planner's objective is to maximize the consumption of a good, and she allocates per-unit co-payments and taxes to its producers. We show that the optimal policy taxes the more efficient firms and allocates larger co-payments to less efficient firms, making it impractical. Therefore, we consider the simple and frequently implemented policy that allocates the same co-payment to each firm, known as *uniform co-payments*, and provide the *first worst-case performance guarantees* for it. Namely, we show that uniform co-payments are guaranteed to induce a significant fraction of the consumption induced by the optimal policy in small markets, for price-taking (resp. Cournot) producers with affine increasing marginal costs facing *any* non-increasing (resp. linear) inverse demand function, even for different firms' efficiency levels. Moreover, when compared to the best policy that allocates co-payments only, uniform co-payments induce *at least half* of the optimal consumption. Furthermore, for Cournot competition with linear demand and constant marginal costs, the latter guarantee increases to *more than 85%* of the optimal consumption. Our results suggest that uniform co-payments are surprisingly powerful in increasing the market consumption of a good.

Key words: Subsidies, Worst-case analysis, Market competition

History:

1. Introduction

We study a class of subsidies and taxes allocation problems where a central planner intervenes in the market of a good by charging taxes or allocating per-unit subsidies, i.e *co-payments*, to its producers, with the objective of increasing its market consumption. The need for this intervention is motivated by the observation that, in many relevant settings, the aggregate market consumption induced by the competition between selfish producers is less than what it is considered socially optimal. This is generally due to the positive societal externalities generated by the consumption of the good. A current example of this issue are the recent efforts around the production of anti-malarial drugs to the developing world (Arrow et al. (2004)). Typically, the central planer decides the co-payments and taxes allocation in the presence of a budget constraint, which is often determined before the allocation decision is made. For example, the role of the central planner could be played by a foundation that has raised money to address the low consumption of an infectious disease treatment, and then faces the problem of deciding how to allocate co-payments and charge taxes subject to this budget. An additional challenge faced by the central planner is that her co-payments and taxes allocation will likely change the market equilibrium attained by the competing producers. Therefore, in order to maximize the market consumption induced by her allocation, the central planer has to take these potentially complex market dynamics into account.

In practice, due to the political difficulty associated with charging new taxes to the producers of a good, many times only co-payments are applied. Moreover, the co-payment allocation policy most often implemented is *uniform*, in the sense that every firm gets the same co-payment regardless of any differences in their cost structure or efficiency. This is probably due to its simplicity, transparency, and ease of implementation, as well as the notion of fairness. In this paper we present the *first worst-case guarantees* for the performance of uniform co-payments in maximizing the market consumption of a good. This provides theoretical support for both continuing with the implementation of uniform co-payments in practice, as well as giving a plausible explanation for the good performance results observed in simulations in previous papers, see Levi et al. (2017). Specifically, our results suggest that the efficiency loss induced by uniform co-payments in maximizing the market consumption of a good can be expected to be relatively small in a large number of settings. Hence, this bounded efficiency loss should be weighed against the important practical advantages of uniform co-payments, such as their political viability, transparency, and ease of implementation and communication.

In order to study these issues, we consider a model that explicitly captures the setting of a central planner aiming at maximizing the market consumption of a good, in the presence of a budget constraint and market competition between heterogeneous profit-maximizing firms. The firms are heterogeneous in terms of their respective marginal cost functions, which model their firm-specific efficiency. In more details, the models studied in this paper capture, as a special case, general non-increasing inverse demand functions, and price-taking producers with affine increasing marginal costs. The generality of an arbitrary non-increasing inverse demand function allows to model complex downstream demand mechanisms. Other special cases included in the class of models we study are important settings with imperfect competition dynamics, such as Cournot competition with linear demand and affine marginal costs.

A practical example of the use of uniform co-payments is the case of the global fight against malaria. In 2004 the Institute of Medicine (IoM) reviewed the economics of the most effective anti-malarial drug available at the time, Artemisinin combination therapies (ACT). It identified that several manufacturers compete in an unregulated market, and concluded that the most effective way of ensuring access to ACTs for the greatest number of patients would be to provide a centralized subsidy to the producers. In this context, the Roll Back Malaria Partnership and the World Bank developed the Affordable Medicines Facility for malaria (AMFm). In 2008, the Global Fund started hosting the AMFm, which began operations in July 2010. By July 2012, the AMFm had managed a budget of US$336 millions -pledged by UNITAID, the governments of the United Kingdom and Canada, and the Bill & Melinda Gates Foundation- to pursue its main objective: increasing the consumption of ACTs, as detailed in the evaluation report by the AMFm Independent Evaluation

Team (2012). Moreover, the AMFm proposed to implement *a uniform co-payment* for the 11 firms participating in the program, which range from large pharmaceuticals like Novartis and Sanofi, to smaller firms in Uganda, India and Korea (for more details see the market intelligence aggregator, funded by UNITAID, A2S2 (2015)). One additional relevant characteristic of the AMFm program is that all the ACT manufacturers that receive co-payments commit to supply anti-malarials on a no profit/no loss basis, see the report by Boulton (2011). Providing the right incentives to the producers can increase access to these drugs, having a significant impact on this global problem (Arrow et al. (2004)).

1.1. Results and Contributions

The main contributions of this paper are the following:

Insights into the structure of optimal co-payments and taxes. We characterize the optimal co-payments and taxes allocation, for price-taking (resp. Cournot) producers with affine increasing marginal costs, facing *any* non-increasing (resp. linear) inverse demand function. We show that the optimal policy *taxes the more efficient firms* and allocates *larger co-payments to less efficient firms*, with the underlying intention of increasing the competition in the market by incorporating additional firms to the market equilibrium that would not join without subsidies, and by exogenously making less efficient firms more competitive and more efficient firms less competitive. However, the structure of the optimal policy generates agency problems, since the firms have an incentive to under-report their efficiency. Moreover, the optimal co-payments and taxes are defined through complicated functions of the firms' marginal cost parameters, which are sensitive to possible parameters' misspecifications. We illustrate the latter observation through numerical simulations. We argue that this policy is hard to implement, supporting the use of the more practical uniform co-payments as long as they induce a small efficiency loss.

Worst-case performance guarantees for uniform co-payments in small markets. We show a *tight* parametric worst-case bound for the performance of uniform co-payments in maximizing market consumption, when compared to the optimal policy that is allowed to charge taxes and allocate co-payments to the firms. This bound is decreasing in the number of firms in the market. In particular, uniform co-payments are guaranteed to induce *at least half* of the optimal market consumption, for markets with a moderate number of firms, specifically nine or less, while the guarantee improves to *at least two-thirds* in markets with four or less firms. This is important because allocating uniform co-payments is significantly less controversial than taxing firms with the goal of increasing the market consumption of a good. On the other hand, we also show that the worst-case performance of *any policy that only allocates co-payments*, when compared to the optimal consumption, deteriorates without bound as the number of firms in the market increases.

Uniform worst-case performance guarantees when only subisides are allowed. When comparing against the best policy that can allocate co-payments only (and not charge taxes), we show that uniform co-payments are guaranteed to induce *at least half* of its market consumption. This result is surprising, particularly since it holds for any number of producers with arbitrarily different levels of efficiency, facing a general non-increasing inverse demand function. Moreover, for the important special case of Cournot competition with linear demand and constant marginal costs, the worst-case performance guarantee for uniform co-payments in maximizing market consumption increases to a surprisingly high 85.36%. Importantly, we show that these worst-case bounds are *asymptotically tight* as the number of firms in the market grows large, hence they cannot be improved. Additionally, for the latter case we obtain even higher *tight* guarantees for any fixed number of firms.

In summary, our results suggest that uniform co-payments are not only simple, but also likely to induce a near-optimal market consumption.

2. Literature Review

In economics, there is a large body of literature that studies the effect of taxes and subsidies. Fullerton and Metcalf (2002) present a thorough review of classical and recent result in this area. This paper is closely related to the study of subsidies in imperfect competition models. However, the traditional approach in this literature assumes homogeneous firms, and focuses on studying the impact of taxes, or subsidies, on the number of firms participating in the market in a symmetric equilibrium, see Fullerton and Metcalf (2002). In contrast, the model in this paper has an operational approach, in the sense the producers are assumed to be heterogeneous, and we focus on the specific subsidy allocation among them, and its dependence on the available budget.

The co-payments allocation problem that we study in this paper was introduced by the authors in Levi et al. (2017), motivated by the practical case of the global subsidy for ACT anti-malarial drugs. Using this framework, sufficient conditions on the firms' marginal cost functions were identified such that uniform co-payments are optimal. Additionally, simulation results in relevant settings where uniform co-payments are not optimal suggested that they are nonetheless very effective. This paper provides a possible explanation for these observations, by taking a worst-case analysis perspective and showing the first performance guarantees for this policy.

The problem of allocating subsidies to increase the market consumption of ACT anti-malarial drugs was studied independently in Taylor and Xiao (2014). They consider the case of one producer selling to multiple heterogeneous retailers facing stochastic demand, and analyze the placement of the subsidy in the supply chain by the central planner. Specifically, they compare subsidizing sales or purchases from the retailers' point of view, concluding that the central planner should only subsidy purchases, which is equivalent to subsidizing the producer in our setting. We consider a different model where we incorporate multiple heterogeneous producers. The procurement and inventory management component of distributing anti-malarials in Africa, and the associated stock-out risk, has been studied through simulations by Gallien et al. (2017). A different, but related, area of research in management science studies the problem of a central planner deciding rebates and subsidies to the *consumers*, with the goal of incentivizing the adoption of green technology or maximizing social welfare, see Aydin and Porteus (2009), Cohen et al. (2016), Krass et al. (2013), and Raz and Ovchinnikov (2015). In contrast, our work is motivated by a different set of practical applications, and it focuses on co-payments that are allocated to the producers.

The special case of Cournot competition with linear demand and constant marginal costs is a basic oligopoly model where the firms compete in quantity. It is a well understood model that provides interesting insights. Therefore, it is frequently used by researchers as a building block to study complex phenomena. Examples of this trend in the operations research and management science literature include using this model, among others, to study the structure of supply chains (Corbett and Karmarkar (2001)), supply chain contracts (Cachon (2003)), production under yield uncertainty (Deo and Corbett (2009)), and firms' profits compared to other equilibrium concepts (Farahat and Perakis (2011)). For this important model, we provide a detailed analysis of the worst-case performance of uniform co-payments in maximizing the market consumption of a good.

3. Model

In this section, we describe a mathematical programming formulation for the problem of allocating co-payment subsidies and charging taxes to the producers of a good, with the objective of maximizing its aggregate consumption while being subject to a budget constraint.

We consider a market for a commodity composed by $n \geq 2$ heterogeneous competing firms. We assume that each firm $i \in \{1, \ldots, n\}$ decides its output $q_i \geq 0$ independently, with the goal of maximizing its own profit. Firms are characterized by an affine increasing marginal cost $h_i(q_i) = c_i + g_i q_i$, where $c_i \geq 0$ and $g_i > 0$. Consumers are described by an inverse demand function $P(Q)$, where $Q = \sum_{i=1}^{n} q_i$ is the aggregate market consumption. We assume that $P(Q)$ is non-negative, non-increasing and continuous in its support $[0, \bar{Q}]$, where $\bar{Q}$ is the smallest value such that $P(\bar{Q}) = 0$, or infinity otherwise. We also assume, without loss of generality, that the firms are labeled such that $0 \leq c_1 \leq c_2 \leq \ldots \leq c_n$, and define $c_0 = 0$.

In terms of the market equilibrium dynamics, we assume that each firm participating in the market equilibrium produces up to the point where its marginal cost equals the market price; and firms that do not participate in the market equilibrium must have a marginal cost of producing their first unit which is at least as large. This can be expressed in the following condition:

$$\text{For each } i, j, \text{ if } \; q_i > 0, \; \text{then} \; c_i + g_i q_i = P(Q) \leq c_j + g_j q_j. \tag{1}$$

The assumptions that the inverse demand function $P(Q)$ is non-increasing, and that the firms' marginal cost functions are increasing, imply the existence and uniqueness of the market equilibrium, see for example Marcotte and Patriksson (2006).

We emphasize that equation (1) corresponds exactly to the Nash equilibrium condition for important imperfect market competition models, such as Cournot competition with linear demand and affine marginal costs. Specifically, a Cournot equilibrium is characterized by, given all the other firms' production levels, each firm i setting its output $q_i \geq 0$ to maximize its profit $\Pi_i = P(Q)q_i - \int_0^{q_i} h_i(x_i)dx_i$. If we assume a linear inverse demand function $P(Q) = a - bQ$, $b \geq 0$, then for any increasing marginal cost functions $h_i(q_i)$ there exists a unique market equilibrium defined by the solution to the first-order conditions of the firms' profit maximization problem, see Vives (2001). Namely, at equilibrium, each firm sets it output at a level such that if $q_i > 0$ then $\frac{\partial \Pi_i}{\partial q_i} = 0$. It is not hard to see that the Cournot equilibrium condition corresponds to equation (1) for modified parameters $\tilde{g}_i = g_i + b$. This equivalence is preserved even when yield uncertainty is incorporated into the model, see Levi et al. (2017). For analytical tractability we focus on linear demand and affine marginal costs, however, it has been demonstrated through numerical simulations that uniform co-payments perform surprisingly well even under Cournot competition with nonlinear demand and nonlinear marginal costs, we refer the reader to Section 7.3 in Levi et al. (2017).

Another special case of condition (1) is the model where the firms act as price-takers and compete in quantity, for any non-increasing inverse demand function. In particular, this model is a reasonable approximation when the firms have little market power, e.g. when there are many firms competing in the market, or when they face the threat of entry to the market, see Tirole (1988). In particular, this simple model captures the behavior of ACT anti-malarial manufacturers discussed in the Introduction, where all the firms receiving co-payments operate in a no-profit/no loss basis.

We focus on settings where the market consumption induced at the market equilibrium is less than what is considered socially optimal. For this reason, a central planner intervenes in the market by allocating a per-unit subsidy -that we refer to as a co-payment- or charging a per-unit tax, to each firm. We refer to the problem faced by the central planner as the *co-payments and taxes allocation problem*, and we denote it by (CTAP). This is a particular case of a Stackelberg game, or a bilevel optimization problem. In the first stage, the central planner decides the co-payments or taxes, denoted by y_i for each firm $i \in \{1,\ldots,n\}$, subject to her budget $B \geq 0$. Specifically, if $y_i \geq 0$ then firm i gets a per-unit subsidy. Similarly, if $y_i < 0$ then firm i is charged a per-unit tax. Moreover, the central planner anticipates that in the second stage the equilibrium output of each firm will satisfy a modified version of the equilibrium condition (1), stated below in equation (5). The main difference between the market equilibrium conditions (5) and (1) is that, from firm i's perspective, the effective price is now $P(Q)+y_i$.

The central planner's objective is to maximize the aggregate market consumption. We refer the reader to Levi et al. (2017) for an extensive motivation of this objective function. Then, the problem faced by the central planner can be formulated as

$$
(CTAP) \qquad \begin{aligned}
\max_{\mathbf{q},Q} \ & Q \\
\text{s.}t. \ & \sum_{i=1}^{n}(c_i + g_i q_i - P(Q))q_i \leq B && (2)\\
& \sum_{j=1}^{n} q_j = Q && (3)\\
& q_i \geq 0, \ \text{for each } i \in \{1,\ldots,n\}. && (4)
\end{aligned}
$$

The co-payments or taxes that the central planner must allocate to induce the outputs $\mathbf{q}$ are

$$y_i = c_i + g_i q_i - P(Q), \ \text{for each firm } i \in \{1,\ldots,n\}. \tag{5}$$

Constraint (2) is the budget constraint, where the total amount spent in co-payments must be at most the available budget B plus the total amount collected in taxes. Equation (5) corresponds to the modified equilibrium condition previously discussed. Note that the fact that we impose the modified market equilibrium condition (5) *on each* firm, does *not* imply that every firm must join the market equilibrium, we refer the reader to Levi et al. (2017) for the details.

Next, we show that we can characterize the market consumption induced (*i*) by optimal co-payments and taxes, (*ii*) by optimal co-payments only (when taxes are not allowed), and (*iii*) by uniform co-payments (when taxes are not allowed and all firms must get the same co-payment).

3.1. Optimal Co-payments and Taxes Allocation

In this section, we characterize the structure of the market equilibrium induced by optimal co-payments and taxes, summarized in Figure 1. Proposition 1 below shows that the market outputs induced by the optimal co-payments and taxes allocation have an intuitive structure. Specifically, more efficient firms have a larger output in equilibrium than less efficient firms, in the sense that if an active firm i dominates firm j in both of the marginal cost function's parameters, namely in its intercept parameter $c_j \geq c_i$ and in its linear parameter $g_j \geq g_i$, then firm i will have a larger market output. Recall that firm i's decision to join the market equilibrium depends on the marginal cost of producing its first unit c_i. This follows from considering its marginal cost function $h_i(q_i) = c_i + g_i q_i$ evaluated at $q_i = 0$. We denote by the index $t \in \{1, \ldots, n\}$ the last firm that has a positive output in this market equilibrium. In other words, any firm with index $i > t$ is so inefficient, in terms of its marginal cost intercept parameter c_i, that it does not join the market equilibrium.

Interestingly, the optimal co-payments and taxes have the following structure. The more efficient firms, in terms of having a smaller value of their marginal cost intercept parameter c_i, may actually get taxed, and this tax is decreasing in c_i. Moreover, only the firms with an index larger than some $s \in \{1, \ldots, t\}$ get a co-payment, and this co-payment is increasing in c_i (see Figure 1 and Proposition 1 below). Namely, in order to maximize the market consumption at equilibrium, for price-taking (resp. Cournot) producers with affine increasing marginal costs facing an arbitrary (resp. linear) non-increasing inverse demand function, the best that the central planner can do is to give *more co-payments to less efficient firms and charge larger taxes to more efficient firms.* This structure suggests that the main driver to maximize the consumption of a good is to increase the competition in the market, by potentially incorporating additional (less efficient) firms into the market equilibrium that would not join without subsidies, and making them more competitive by means of larger co-payments. Moreover, such co-payments can be partially funded by taxing the more efficient firms, making the latter less competitive. Note that the difference between the optimal co-payments and taxes for different firms in equation (6) depends on the difference between

$$0 \le \overbrace{c_1 \le \ldots \le c_{s-1} \le c_s \le \ldots \le c_t}^{\text{Active firms: } q_i^{\mathrm{CT}} \ge 0} \le \overbrace{c_{t+1} \le \ldots \le c_n}^{\text{Inactive firms: } q_i^{\mathrm{CT}} = 0}$$

$$\underbrace{y_1^{\mathrm{CT}} \le \ldots \le y_{s-1}^{\mathrm{CT}}}_{\text{Decreasing taxes}} < 0 \le \underbrace{y_s^{\mathrm{CT}} \le \ldots \le y_t^{\mathrm{CT}}}_{\text{Increasing co-payments}}, \underbrace{y_{t+1}^{\mathrm{CT}} = \ldots = y_n^{\mathrm{CT}} = 0}_{\text{No co-payments}}$$

Figure 1 Market equilibrium induced by optimal co-payments and taxes

their marginal cost intercept parameters c_i, i.e. on how much firms need to be incentivized to join the market equilibrium. However, we emphasize that the absolute value of the optimal co-payments does depend on both marginal cost parameters c_i and g_i, see equation (8).

PROPOSITION 1. *Any optimal solution to problem (CTAP),* $(\boldsymbol{q}^{\mathrm{CT}}, Q^{\mathrm{CT}})$*, must be such that the budget constraint (2) is tight, and there exists an index* $t \in \{1, \ldots, n\}$*, such that the optimal co-payments and taxes are given by*

$$y_i^{\mathrm{CT}} = y_1^{\mathrm{CT}} + \frac{c_i - c_1}{2} > 0 \textit{ for each } i \in \{1, \ldots, t\}, \quad y_i^{\mathrm{CT}} = 0 \textit{ for each } i \in \{t+1, \ldots, n\}. \tag{6}$$

Moreover, there exists an index $s \in \{1, \ldots, t\}$*, such that* $y_i^{\mathrm{CT}} < 0$ *for all* $i < s$*, and* $y_i^{\mathrm{CT}} \ge 0$ *for all* $i \ge s$*. The optimal market outputs are given by*

$$q_1^{\mathrm{CT}} = \frac{P(Q^{\mathrm{CT}}) + y_1^{\mathrm{CT}} - c_1}{g_1}, \; q_i^{\mathrm{CT}} = \frac{g_1}{g_i} q_1^{\mathrm{CT}} - \frac{c_i - c_1}{2g_i} > 0 \textit{ for each } i \in \{1, \ldots, t\},$$
$$q_i^{\mathrm{C}} = 0 \textit{ for each } i \in \{t+1, \ldots, n\}. \tag{7}$$

The expressions (6)-(7) are written as a function of y_1^{CT}*, which is such that*

$$\sum_{i=1}^{t} \frac{y_1^{\mathrm{CT}}}{g_i} = \sqrt{\left(\sum_{i=1}^{t} \frac{P(Q^{\mathrm{CT}}) - c_i}{2g_i}\right)^2 + \sum_{i=1}^{t} \frac{B}{g_i} + \sum_{i=1}^{t} \frac{c_i^2}{2g_i} \sum_{i=1}^{t} \frac{1}{2g_i} - \left(\sum_{i=1}^{t} \frac{c_i}{2g_i}\right)^2} - \sum_{i=1}^{t} \frac{P(Q^{\mathrm{CT}}) - c_1}{2g_i}. \tag{8}$$

Finally, the aggregate market consumption Q^{CT} *must satisfy the following fixed point equation*

$$Q^{\mathrm{CT}} = \sqrt{\left(\sum_{i=1}^{t} \frac{P(Q^{\mathrm{CT}}) - c_i}{2g_i}\right)^2 + \sum_{i=1}^{t} \frac{B}{g_i} + \sum_{i=1}^{t} \frac{c_i^2}{2g_i} \sum_{i=1}^{t} \frac{1}{2g_i} - \left(\sum_{i=1}^{t} \frac{c_i}{2g_i}\right)^2} + \sum_{i=1}^{t} \frac{P(Q^{\mathrm{CT}}) - c_i}{2g_i}. \tag{9}$$

The proof of Proposition 1 follows from the KKT conditions of problem (CTAP), which can be shown to be necessary for optimality, and it is omitted for the sake of brevity.

3.2. Optimal Co-payments Allocation

In this section, we additionally assume that the central planner can only allocate co-payments and never tax the producers. In other words, the central planner's decision variables y_i must be non-negative. This is the case in most practical applications of this problem. Specifically, the central planner usually has no authority to tax firms that may operate in different countries, with the goal of increasing the aggregate market consumption of a good, see for example Arrow et al. (2004) for the case of anti-malarials. Formally, we add the constraint

$$c_i + g_i q_i \geq P(Q), \text{ for each } i \in \{1, \dots, n\} \tag{10}$$

to problem (CTAP), we refer to the resulting formulation as the *co-payment allocation problem*, denoted by (CAP). From equation (5) it follows that constraint (10) is equivalent to $y_i \geq 0$ for each firm $i \in \{1, \dots, n\}$.

We now characterize the structure of the market equilibrium induced by optimal co-payments, summarized in Figure 2. Proposition 2 below shows that the market outputs induced by the optimal co-payments allocation, and the optimal co-payments themselves, preserve the structure of the optimal co-payments and taxes discussed in Section 3.1, except that constraint (10) plays an active role in preventing the taxation of the more efficient firms. Specifically, as before, more efficient firms have a larger output at equilibrium than less efficient firms. We denote by the index $m \in \{1, \dots, n\}$ the last firm that has a positive output in this market equilibrium. Additionally, the more efficient firms, in terms of having a smaller value of their marginal cost intercept parameter c_i, may get no co-payments. Only after some index $l \in \{1, \dots, m\}$, the firms start getting a co-payment that is increasing in c_i (see Figure 2 and Proposition 2 below). Namely, to maximize the market consumption in this model, the best that the central planner can do using co-payments only is to give *more co-payments to less efficient firms.*

Proposition 2. *Any optimal solution of the co-payments allocation problem (CAP),* $(\boldsymbol{q}^{\mathrm{C}}, Q^{\mathrm{C}})$*, must be such that the budget constraint (2) is tight, and there exist indexes* $l, m \in \{1, \dots, n\}$*, with* $l \leq m$*, such that the optimal co-payments are given by*

$$y_i^{\mathrm{C}} = 0 \textit{ for each } i \in \{1, \dots, l-1\}, \; y_i^{\mathrm{C}} = y_l^{\mathrm{C}} + \frac{c_i - c_l}{2} > 0 \textit{ for each } i \in \{l, \dots, m\},$$

$$0 \leq \overbrace{c_1 \leq \ldots \leq c_{l-1} \leq c_l \leq \ldots \leq c_m}^{\text{Active firms: } q_i^{\mathrm{C}} \geq 0} \leq \overbrace{c_{m+1} \leq \ldots \leq c_n}^{\text{Inactive firms: } q_i^{\mathrm{C}} = 0}$$

$$\underbrace{0 = y_1^{\mathrm{C}} = \ldots = y_{l-1}^{\mathrm{C}}}_{\text{No co-payments}} < \underbrace{y_l^{\mathrm{C}} \leq \ldots \leq y_m^{\mathrm{C}}}_{\text{Increasing co-payments}}, \underbrace{y_{m+1}^{\mathrm{C}} = \ldots = y_n^{\mathrm{C}} = 0}_{\text{No co-payments}}$$

Figure 2 Market equilibrium induced by optimal co-payments

$$y_i^{\mathrm{C}} = 0 \textit{ for each } i \in \{m+1, \ldots, n\}. \tag{11}$$

The optimal market outputs are given by

$$q_i^{\mathrm{C}} = \frac{g_l}{g_i} q_l^{\mathrm{C}} + \frac{c_l - c_i}{g_i} - \frac{y_l^{\mathrm{C}}}{g_i} \geq \frac{g_l}{g_i} q_l^{\mathrm{C}} + \frac{c_l - c_i}{2g_i} > 0 \textit{ for each } i \in \{1, \ldots, l-1\},\ q_l^{\mathrm{C}} = \frac{P(Q^{\mathrm{C}}) + y_l^{\mathrm{C}} - c_l}{g_l}, \tag{12}$$

$$q_i^{\mathrm{C}} = \frac{g_l}{g_i} q_l^{\mathrm{C}} - \frac{c_i - c_l}{2g_i} > 0 \textit{ for each } i \in \{l, \ldots, m\},\ q_i^{\mathrm{C}} = 0,\ \textit{for each } i \in \{m+1, \ldots, n\}. \tag{13}$$

The expressions (11)-(13) are written as a function of the first positive co-payment y_l^{C}, which is such that

$$\sum_{i=l}^{m} \frac{y_l^{\mathrm{C}}}{g_i} = \sqrt{\left(\sum_{i=l}^{m} \frac{P(Q^{\mathrm{C}}) - c_i}{2g_i}\right)^2 + \sum_{i=l}^{m} \frac{B}{g_i} + \sum_{i=l}^{m} \frac{1}{g_i} \left(\sum_{i=l}^{m} \left(\frac{c_i - c_l}{2}\right)^2 \frac{1}{g_i}\right) - \left(\sum_{i=l}^{m} \frac{c_i - c_l}{2g_i}\right)^2} - \sum_{i=l}^{m} \frac{P(Q^{\mathrm{C}}) - c_l}{2g_i}. \tag{14}$$

Finally, the aggregate market consumption Q^{C} must satisfy the following fixed point equation

$$\begin{aligned} Q^{\mathrm{C}} = & \sqrt{\left(\sum_{i=l}^{m} \frac{P(Q^{\mathrm{C}}) - c_i}{2g_i}\right)^2 + \sum_{i=l}^{m} \frac{B}{g_i} + \sum_{i=l}^{m} \frac{1}{g_i} \left(\sum_{i=l}^{m} \left(\frac{c_i - c_l}{2}\right)^2 \frac{1}{g_i}\right) - \left(\sum_{i=l}^{m} \frac{c_i - c_l}{2g_i}\right)^2} \\ & + \sum_{i=1}^{m} \frac{P(Q^{\mathrm{C}}) - c_i}{g_i} - \sum_{i=l}^{m} \frac{P(Q^{\mathrm{C}}) - c_i}{2g_i}. \end{aligned} \tag{15}$$

The proof of Proposition 2 follows from the KKT conditions of problem (CAP), which can be shown to be necessary for optimality, and it is omitted for the sake of brevity.

3.3. Uniform Co-payments

In this section, we characterize the structure of the market equilibrium induced by uniform co-payments, summarized in Figure 3. By definition, this policy allocates the same co-payment to each

$$0 \leq \overbrace{c_1 \leq \ldots \leq c_u}^{\text{Active firms: } q_i^{\mathrm{U}} \geq 0} \leq \overbrace{c_{u+1} \leq \ldots \leq c_n}^{\text{Inactive firms: } q_i^{\mathrm{U}} = 0}$$

$$\underbrace{\frac{B}{Q^{\mathrm{U}}} = y_1^{\mathrm{U}} = \ldots = y_u^{\mathrm{U}}}_{\text{Uniform co-payments}}, \underbrace{y_{u+1}^{\mathrm{U}} = \ldots = y_n^{\mathrm{U}} = 0}_{\text{No co-payments}}$$

Figure 3 Market equilibrium induced by uniform co-payments

firm with a positive output. However, highly inefficient firms might not produce at all and hence receive no co-payment in the induced market equilibrium. We denote by the index $u \in \{1, \ldots, n\}$ the last firm that has a positive output in the market equilibrium induced by uniform co-payments.

Naturally, we focus on the largest possible uniform co-payment that can be afforded with the central planner's budget B. Specifically, if we denote by q_i^{U} firm i's output induced by the uniform co-payment y^{U}, and by Q^{U} the total market output, then we focus on y^{U} such that $\sum_{i=1}^{n} q_i^{\mathrm{U}} y^{\mathrm{U}} = B$, or equivalently $y^{\mathrm{U}} = \frac{B}{Q^{\mathrm{U}}}$. Namely, the amount of the uniform co-payment is obtained by simply dividing the available budget B by the largest market consumption that be can attained with this budget under a uniform co-payment policy, denoted by Q^{U}. In practice, the way this policy is usually implemented is by dividing the budget by a target market consumption that the central planner has set as a goal, see for example AMFm Independent Evaluation Team (2012) for the case of ACT anti-malarials. In terms of the parameters in our model, the structure of the market equilibrium induced by uniform co-payments is described in the following proposition.

PROPOSITION 3. *Assume that the inverse demand function $P(Q)$ is non-negative, non-increasing, and continuous in $[0, \bar{Q}]$, and that the affine marginal cost function of each firm $i \in \{1, \ldots, n\}$ is $h_i(q_i) = c_i + g_i q_i$, where $c_i \geq 0$ and $g_i > 0$. Then, the market consumption induced by the uniform co-payments allocation of the budget $B \geq 0$, Q^{U}, is the unique solution to the following fixed point equation*

$$Q^{\mathrm{U}} = \sum_{i=1}^{u} \frac{P(Q^{\mathrm{U}}) - c_i}{2g_i} + \sqrt{\left(\sum_{i=1}^{u} \frac{P(Q^{\mathrm{U}}) - c_i}{2g_i}\right)^2 + \sum_{i=1}^{u} \frac{B}{g_i}}, \tag{16}$$

where the index $u \in \{1,\dots,n\}$ *is such that* $\sum_{i=1}^{u} \frac{c_u - c_i}{g_i} \le Q^{\mathrm{U}} \le \sum_{i=1}^{u} \frac{c_{u+1}-c_i}{g_i}$*, and we define, without loss of generality,* $c_{n+1} = P(0)$*. Similarly, the firms' outputs* q_i^{U} *are*

$$q_i^{\mathrm{U}} = \frac{Q^{\mathrm{U}} + \sum_{j=1}^{u} \frac{c_j}{g_j}}{g_i \sum_{j=1}^{u} \frac{1}{g_j}} - \frac{c_i}{gi} \ge 0 \quad \textit{for each } i \in \{1,\dots,u\}, \quad q_i^{\mathrm{U}} = 0, \textit{ for each } i \in \{u+1,\dots,n\}. \tag{17}$$

The uniform co-payment that induces this market output is

$$y_i^{\mathrm{U}} = \frac{B}{Q^{\mathrm{U}}} \ge 0 \quad \textit{for each } i \in \{1,\dots,u\}, \quad y_i^{\mathrm{U}} = 0, \textit{ for each } i \in \{u+1,\dots,n\}. \tag{18}$$

The proof of Proposition 3 follows directly from the market equilibrium condition (5) together with the fact that each active firm receives the same co-payment $y_i^{\mathrm{U}} = \frac{B}{Q^{\mathrm{U}}}$, and it is omitted.

3.4. Policy Comparison

In this section, we compare the three policies we have defined in the previous subsections. Namely, we compare the market equilibrium induced by the optimal co-payments and taxes allocation $(\mathbf{q}^{\mathrm{CT}}, Q^{\mathrm{CT}}, \mathbf{y}^{\mathrm{CT}})$, to the one induced by optimal co-payments $(\mathbf{q}^{\mathrm{C}}, Q^{\mathrm{C}}, \mathbf{y}^{\mathrm{C}})$, and uniform co-payments $(\mathbf{q}^{\mathrm{U}}, Q^{\mathrm{U}}, \mathbf{y}^{\mathrm{U}})$, respectively, along several dimensions.

Implementation Challenges. Both $\mathbf{y}^{\mathrm{CT}}$ and $\mathbf{y}^{\mathrm{C}}$ are defined through complicated functions of the problem parameters, which makes them difficult to communicate. Moreover, they are different for each firm, which would significantly increase the complexity of paying subsidies and charging taxes to the producers. In contrast, the uniform co-payments policy is simple to communicate and control. For example, the value of the uniform co-payment can be effectively decided in a meeting. In fact, as mentioned before, the uniform co-payment is usually implemented by dividing the budget by a target market consumption that the central planner has set as a goal. Additionally, any implementation of the optimal co-payment and taxes allocation $\mathbf{y}^{\mathrm{CT}}$, would have to deal with the controversial issue of taxing the more efficient firms, and subsidizing the less efficient firms. Furthermore, taxing firms may simply not be a realistic possibility if the central planner role is played by an NGO, or an international institution such as the World Bank.

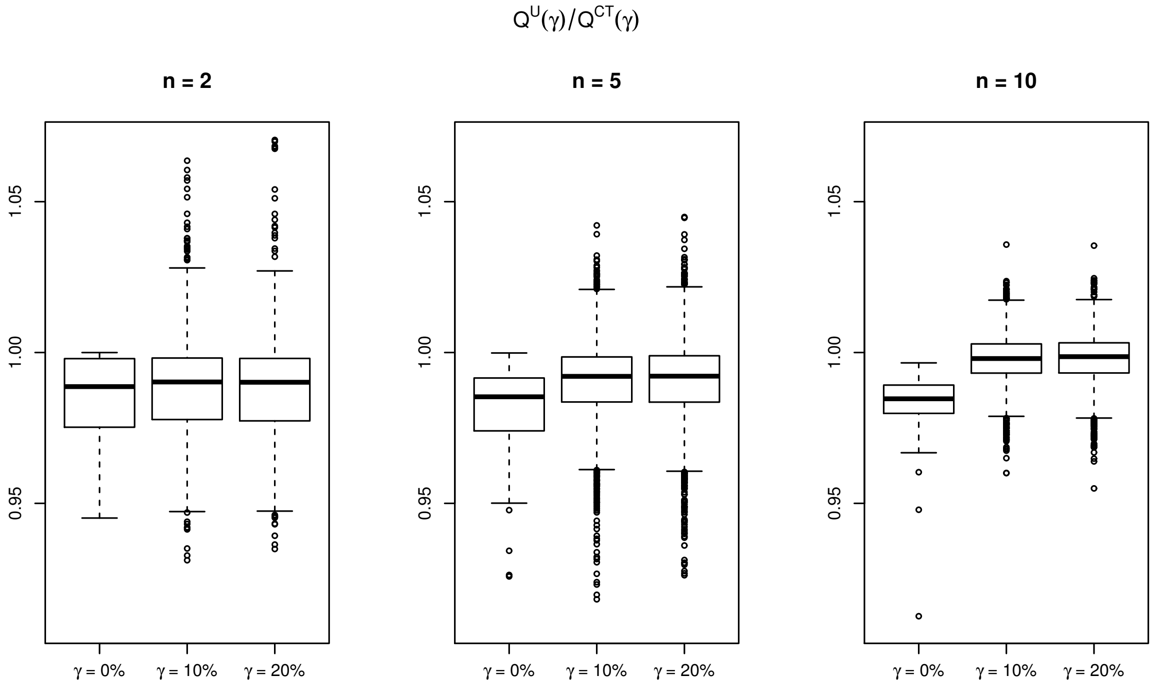

Figure 4 Relative performance of uniform co-payments for Cournot competition with linear demand and *uncertain* constant marginal costs, parametrized by the number of firms n and the uncertainty level γ.

Information Requirements. Both $\mathbf{y}^{\mathrm{CT}}$ and $\mathbf{y}^{\mathrm{C}}$ require the central planner to know *each individual* marginal cost function $h_i(q_i) = c_i + g_i q_i$. Although we assumed a full information setting, this may not be the case in practice. Therefore, any practical implementation of $\mathbf{y}^{\mathrm{C}}$ (or $\mathbf{y}^{\mathrm{CT}}$) would require a truthful mechanism to elicit the marginal cost functions, or alternatively it would have to deal with potential misspecifications. This is further complicated by the agency problems induced by them. Namely, if less efficient firms receive larger co-payments, then firms have a clear incentive to misrepresent their marginal costs. In contrast, the uniform co-payment depends on *averages* of simple functions of the marginal costs' parameters of the active firms in the market equilibrium, as well as the inverse demand function and the available budget. Specifically, it only depends on $\sum_{i=1}^{u} \frac{1}{g_i}$, $\sum_{i=1}^{u} \frac{c_i}{g_i}$, $P(Q)$ and B, see (16)-(18).

Robustness to Data Errors. The dependence of the uniform co-payment on *averages* of the marginal costs' parameters makes Q^{U} more robust to misspecifications on their values, when compared to Q^{CT} or Q^{C}. We demonstrate this feature by simulation. In particular, Figure 4 and Table 1

$Q^U(\gamma)/Q^{CT}(\gamma)$	n = 2			n = 5			n = 10		
	$\gamma = 0$	$\gamma = .1$	$\gamma = .2$	$\gamma = 0$	$\gamma = .1$	$\gamma = .2$	$\gamma = 0$	$\gamma = .1$	$\gamma = .2$
Min.	0.9451	0.9311	0.9349	0.9258	0.9182	0.9262	0.9126	0.9600	0.9550
1st Qu.	0.9753	0.9778	0.9774	0.9741	0.9836	0.9836	0.9799	0.9932	0.9932
Median	0.9887	0.9902	0.9901	0.9808	0.9921	0.9922	0.9847	0.9980	0.9986
Mean	0.9841	0.9869	0.9867	0.9808	0.9909	0.9909	0.9831	0.9977	0.9980
3rd Qu.	0.9979	0.9982	0.9980	0.9914	0.9985	0.9989	0.9892	1.0029	1.0032
Max.	1.0000	1.0636	1.0704	0.9998	1.0421	1.0449	0.9966	1.0358	1.0354
# Samples	100	3000	3000	100	3000	3000	100	3000	3000

Table 1 Relative performance of uniform co-payments for Cournot competition with linear demand and *uncertain* constant marginal costs, parametrized by the number of firms n and the uncertainty level γ.

illustrate the robustness of uniform co-payments for Cournot competition with linear demand and *uncertain* constant marginal costs. Specifically, if the true marginal costs are within γ of the central planner's estimates ($\gamma \in \{10\%, 20\%\}$), then the relative performance of uniform co-payments *improves* on average. Moreover, the market consumption induced by uniform co-payments under the true marginal costs, $Q^U(\gamma)$, many times *outperforms* the one induced by the optimal co-payment and taxes policy, $Q^{CT}(\gamma)$, particularly as the number of firms increases. Comparing against the market consumption induced by optimal co-payments, $Q^C(\gamma)$, leads to similar results. We briefly describe the experimental setup, the details are provided in the Appendix. The number of firms is $n \in \{2, 5, 10\}$. For each n, we generate 100 instances uniformly at random. For each true marginal cost c_i, we sample 30 central planner's estimates $\tilde{c}_i$ from the support $[c_i(1-\gamma), c_i(1+\gamma)]$, for each $\gamma \in \{0.1, 0.2\}$. For each of the 3,000 central planner's estimates $\tilde{\mathbf{c}}$, we compute the optimal co-payments and taxes $\mathbf{y}^{CT}(\tilde{\mathbf{c}})$, optimal co-payments $\mathbf{y}^{C}(\tilde{\mathbf{c}})$, and uniform co-payments $\mathbf{y}^{U}(\tilde{\mathbf{c}})$. Finally, the associated true marginal cost $\mathbf{c}$ is used to compute the induced true market consumptions $Q^{CT}(\gamma)$, $Q^C(\gamma)$, and $Q^U(\gamma)$. Figure 4 and Table 1 then report the boxplots of $Q^U(\gamma)/Q^{CT}(\gamma)$ for each pair (n, γ). We emphasize that when the central planner's estimates are correct (i.e. $\gamma = 0$), the relative performance of uniform co-payments matches the experimental results in Levi et al. (2017).

Naturally, without uncertainty $Q^U(0)/Q^{CT}(0)$ is upper-bounded by 1. However, the main insight from Figure 4 and Table 1 is that even a small relative error on the central planner's estimates can be enough for uniform co-payments to outperform the optimal policy in many cases.

Overall, these characteristics make uniform co-payments a more attractive policy in practice, as long as its loss with respect to the induced market consumption is not very large.

4. Uniform Co-payments versus Optimal Co-payments and Taxes

In this section, we study the worst-case performance of uniform co-payments in maximizing the market consumption of a good, when compared to the optimal policy that can both allocate co-payments and charge taxes to the firms.

Interestingly, we show that uniform co-payments are guaranteed to induce *at least half* of the market consumption induced by the optimal policy, for markets with a moderate number of firms, specifically nine or less. This is an important insight because allocating uniform co-payments is significantly less controversial than taxing some firms in order to increase the market consumption of a good. Moreover, as discussed in Section 3.4 the central planner may not have the authority to charge taxes to the firms, e.g. if this role is played by an NGO or an international institution. Hence, identifying conditions such that the efficiency loss induced by uniform co-payments is bounded, when comparing it to the stringent benchmark of potentially taxing some firms, provides strong theoretical support to continue their use in practice. On the other hand, we also show that the worst-case performance in maximizing market consumption of *any policy that allocates only co-payments* can be arbitrarily bad, when compared to the optimal policy that is additionally allowed to charge taxes, as the number of firms in the market grows very large.

Theorem 1 below provides the main insight in this section.

THEOREM 1. *For any number $n \geq 2$ of price-taking (resp. Cournot) producers with affine increasing marginal cost $h(q_i) = c_i + g_i q_i$ for each firm $i \in \{1, \dots, n\}$, where $0 \leq c_1 \leq \dots \leq c_n$ and $0 < g_1 \leq \dots \leq g_n$, facing a general (resp. linear) non-increasing inverse demand function $P(Q)$, where $\sum_{i=1}^{n} q_i = Q$, and for any budget $B \geq 0$, let Q^{CT} be the market consumption induced by the optimal policy in problem (CTAP), and Q^{U} the consumption induced by uniform co-payments. Then $\frac{Q^{\mathrm{U}}}{Q^{\mathrm{CT}}} \geq \frac{2}{1+\sqrt{n}}$.*

Proof. For any $n \geq 2$, and for any instance of problem (CTAP) such that $h(q_i) = c_i + g_i q_i$ for each firm $i \in \{1,\ldots,n\}$, where $0 \leq c_1 \leq \ldots \leq c_n$ and $0 < g_1 \leq \ldots \leq g_n$, let $(\mathbf{q}^{\mathrm{CT}}, Q^{\mathrm{CT}})$ and $(\mathbf{q}^{\mathrm{U}}, Q^{\mathrm{U}})$ be as defined in Propositions 1 and 3, respectively.

From Proposition EC.1 in the e-companion to this paper it follows that, in order to study the worst-case performance of uniform co-payments, there is no loss of generality in focusing on instances of problem (CTAP) such that $t = u = n$, i.e. on instances where all the firms participate in both the optimal market equilibrium and the one induced by uniform co-payments. We assume this in the rest of the proof.

Note that then the fixed point equation (9) can be re-written as

$$Q^{\mathrm{CT}} = \sum_{i=1}^{n} \frac{1}{g_i}\left(P(Q^{\mathrm{CT}}) + \frac{B}{Q^{\mathrm{CT}}}\right) - \sum_{i=1}^{n} \frac{c_i}{g_i} + \frac{1}{4Q^{\mathrm{CT}}}\left(\sum_{i=1}^{n} \frac{c_i^2}{g_i} \sum_{i=1}^{n} \frac{1}{g_i} - \left(\sum_{i=1}^{n} \frac{c_i}{g_i}\right)^2\right). \tag{19}$$

Similarly, the fixed point equation (16) can be re-written as

$$Q^{\mathrm{U}} = \sum_{i=1}^{n} \frac{1}{g_i}\left(P(Q^{\mathrm{U}}) + \frac{B}{Q^{\mathrm{U}}}\right) - \sum_{i=1}^{n} \frac{c_i}{g_i}. \tag{20}$$

To conclude, note that

$$Q^{\mathrm{CT}}(Q^{\mathrm{CT}} - Q^{\mathrm{U}}) \leq \frac{1}{4}\left(\sum_{i=1}^{n} \frac{c_i^2}{g_i} \sum_{i=1}^{n} \frac{1}{g_i} - \left(\sum_{i=1}^{n} \frac{c_i}{g_i}\right)^2\right) \leq \frac{(n-1)}{4}\left(\sum_{i=1}^{n} \frac{c_n - c_i}{g_i}\right)^2 \leq \frac{(n-1)}{4}(Q^{\mathrm{U}})^2. \tag{21}$$

Where the first inequality follows from equations (19), (20) and $Q^{\mathrm{CT}} \geq Q^{\mathrm{U}}$, the second inequality follows from Lemma EC.1 in the e-companion to this paper, where the non-crossing assumption made on the marginal cost parameters in the statement of the theorem is used, and the third inequality follows from Proposition 3 for $u = n$. Note that (21) is equivalent to $\frac{Q^{\mathrm{CT}}}{Q^{\mathrm{U}}}\left(\frac{Q^{\mathrm{CT}}}{Q^{\mathrm{U}}} - 1\right) \leq \frac{(n-1)}{4}$, which implies the inequality in the Theorem for any non-negative value of $\frac{Q^{\mathrm{CT}}}{Q^{\mathrm{U}}}$. $\square$

Theorem 1 makes the assumption that the marginal costs of the firms have a non-crossing property. In particular, this assumption is satisfied by Cournot competition with linear demand and constant marginal costs. Theorem 1 suggests that if the number of firms in the market is moderate, then the simple policy of allocating the same co-payment to each firm is guaranteed to induce a significant fraction of the optimal market consumption, even if the problem is faced by

	Optimal Co-payments & Taxes	**Co-payments Only**
q_1	$\frac{1}{2}+\frac{1}{2\sqrt{n}}$	1
$q_i,\ i\in\{2,\ldots,n\}$	$\frac{1}{2\sqrt{n}}$	0
Q	$\frac{1+\sqrt{n}}{2}$	1
y_1	$\frac{-(n-1)cq_n^{\mathrm{CT}}}{1+\sqrt{n}}$	0
$y_i,\ i\in\{2,\ldots,n\}$	$\frac{cq_1^{\mathrm{CT}}}{1+\sqrt{n}}$	0

Table 2 Market Equilibria Induced in the Worst-Case Instance

a government that could potentially also charge taxes. Specifically, the worst-case guarantee for uniform co-payments in Theorem 1 evaluates to one-half if the number of firms in the market is nine or less, and to two-thirds if the number of firms in the market is four or less.

Importantly, Proposition 4 below asserts that the bound in Theorem 1 cannot be improved. Additionally, Proposition 4 shows that the non-crossing assumption on the marginal cost functions made in Theorem 1 is *necessary* in order to get a parametric worst-case guarantee for the performance of any policy that allocates co-payments only.

Proposition 4. *The bound in Theorem 1 is tight. Namely, for any number of firms* $n\geq 2$*, there exists an instance of problem (CTAP) such that if we let* Q^{CT} *be the market consumption induced by the optimal co-payments and taxes allocation, and* Q^{C} *be the market consumption induced by the optimal co-payments policy then* $\frac{Q^{\mathrm{C}}}{Q^{\mathrm{CT}}}=\frac{2}{1+\sqrt{n}}$.

Moreover, if the marginal cost functions are allowed to cross, then there exists an instance with $n=2$ *firms where the ratio* $\frac{Q^{\mathrm{C}}}{Q^{\mathrm{CT}}}$ *can be made arbitrarily small.*

The proof of Proposition 4 is presented in the e-companion to this paper. In order to prove the first statement, we exhibit an instance of problem (CTAP) with one efficient firm and $(n-1)$ homogeneous inefficient firms, where the central planner has no budget, i.e. $B=0$, and the market equilibrium induced by the optimal co-payments and taxes allocation, and by *any* policy that allocates co-payments only (in particular the optimal one), are summarized in Table 2. The fact that the worst-case instance has no budget is intuitive, since in this case any policy that is restricted to allocating only co-payments has no impact on the market consumption. On the other hand, a policy that is allowed to tax the more efficient firms can collect a positive budget and use it to subsidize the less efficient firms. In order to prove the second statement in the proposition, we consider a modification of this instance where the marginal costs are allowed to cross and there are only two firms in the market.

Proposition 4 asserts that if the central planner is allowed to tax the firms, then the performance of any policy that allocates only co-payments can be arbitrarily bad as the number of firms grows very large. Moreover, from the second statement in the proposition, if the non-crossing assumption on the marginal cost functions is relaxed then the performance of any policy that allocates co-payments only can be arbitrarily bad, even in a small market with two firms. This result suggests that implementing any policy that allocates co-payments only may not be such a good idea if the central planner is in the position of taxing firms. Having said that, the instances in the proof of Proposition 4 are somewhat degenerate in that the same performance is attained by a policy that does not intervene the market. The latter will change in the next Section, where we will have a distinct worst-case performance between optimal and uniform co-payments, and between them and not intervening in the market.

5. Uniform Co-payments versus Optimal Co-payments Only

In this section, we study the worst-case performance of uniform co-payments when compared to the optimal co-payments policy, which is allowed to allocate firm-dependent co-payments (but not allowed to charge taxes). This is important, since in most practical cases the central planner's role is played by an NGO, or by the World Bank, while the producers are firms operating in different countries (see for example AMFm Independent Evaluation Team (2012) for the case of anti-malarials), therefore charging taxes to the firms may not be a tool available for the central planner. In these cases, our results show that the efficiency loss incurred by implementing the simple and practical policy of uniform co-payments is bounded, and not too large.

Additionally, from Propositions 2 and 3 in Section 3 it is not hard to see that there exist instances where the aggregate market consumption induced by optimal co-payments Q^{C}, and by uniform co-payments Q^{U}, are strictly increasing in the budget available to the central planner B. Therefore, the worst-case performance of the policy that does not intervene the market can be arbitrarily bad when compared to both optimal and uniform co-payments. For this reason, we restrict our attention to uniform co-payments in the rest of this section.

5.1. General Worst-case Performance Guarantee of One Half

In this section, we show an *asymptotically tight* worst-case guarantee of one half for the performance of uniform co-payments in maximizing the market consumption of a good. We start by formally stating the main result.

THEOREM 2. *For any number $n \geq 2$ of price-taking (resp. Cournot) producers with affine increasing marginal cost $h(q_i) = c_i + g_i q_i$, where $c_i \geq 0$ and $g_i > 0$ for each firm $i \in \{1, \ldots, n\}$, facing a non-negative, non-increasing, and continuous (resp. linear) inverse demand function $P(Q)$, where $\sum_{i=1}^{n} q_i = Q$, and for any budget $B \geq 0$, let Q^{C} be the market consumption induced by optimal co-payments, and Q^{U} be the market consumption induced by uniform co-payments. Then $2Q^{\mathrm{U}} \geq Q^{\mathrm{C}}$.*

Theorem 2 asserts that uniform co-payments, despite their simplicity, are guaranteed to induce at least half of the market consumption induced by optimal co-payments in this model. This result is surprising, considering that this worst-case bound holds for *any* number of price-taking (resp. Cournot) producers with arbitrarily different affine increasing marginal costs, facing *any* non-increasing (resp. linear) inverse demand function. In particular, the marginal cost functions are allowed to cross (as opposed to Theorem 1). Theorem 2 provides theoretical support to continue the use of this simple policy in practice, by showing that the potential gain obtained from implementing more complex firm-dependent co-payment allocation policies is limited.

We provide here a proof sketch for Theorem 2, while the full analysis is done in the e-companion to this paper.

Proof Sketch. We use the characterizations of the market equilibria induced by optimal and uniform co-payments to write a mathematical program that minimizes the difference $2Q^{\mathrm{U}} - Q^{\mathrm{C}}$, in terms of the parameters that define a valid instance of the co-payment allocation problem (CAP). Namely, in terms of the firms' marginal cost parameters c_i and g_i, the budget available to the central planner B, and the inverse demand function $P()$. We then show that its optimal objective value is *non-negative*, hence the bound in Theorem 2 holds. Specifically, we show that the optimal objective value of the following mathematical program, denoted by (P^n_{lmu}), is non-negative, for any number of firms n, and fixed values of the indexes l, m, u, described in Figures 2 and 3.

$$
\begin{aligned}
&\min_{\mathbf{c},\mathbf{g},B,P()} && 2Q^{\mathrm{U}} - Q^{\mathrm{C}}_l \\
&\text{s.}t. && y^{\mathrm{C}}_i = 0 \text{ for each } i \in \{1,\ldots,l-1\}, y^{\mathrm{C}}_i \geq 0 \text{ for each } i \in \{l,\ldots,m\} && (22)\\
&(P^n_{lmu}) && q^{\mathrm{C}}_i \geq 0 \text{ for each } i \in \{1,\ldots,m\}, q^{\mathrm{C}}_i = 0 \text{ for each } i \in \{m+1,\ldots,n\} && (23)\\
& && q^{\mathrm{U}}_i \geq 0 \text{ for each } i \in \{1,\ldots,u\}, q^{\mathrm{U}}_i = 0 \text{ for each } i \in \{u+1,\ldots,n\} && (24)\\
& && B \geq 0 && (25)\\
& && c_i \geq c_{i-1}, \text{ for each } i \in \{1,\ldots,n\} && (26)\\
& && g_i \geq 0, \text{ for each } i \in \{1,\ldots,n\} && (27)\\
& && P'(Q) \leq 0, P(Q) \geq 0, \text{ for each } Q \geq 0. && (28)
\end{aligned}
$$

It should be clear we use the notation $(\mathbf{q}^{\mathrm{C}}, Q^{\mathrm{C}}, \mathbf{y}^{\mathrm{C}})$ and $(\mathbf{q}^{\mathrm{U}}, Q^{\mathrm{U}}, \mathbf{y}^{\mathrm{U}})$ to denote the closed-form expressions provided in Propositions 2 and 3, respectively, which are a function of $(\mathbf{c}, \mathbf{g}, B, P())$. For the sake of clarity, we dropped the explicit dependence on $(\mathbf{c}, \mathbf{g}, B, P())$ from the notation. Then, constraints (22)-(24) refer to the feasibility of the market equilibria induced by optimal co-payments and uniform co-payments, respectively. While constraints (25)-(28) refer to the assumptions made on the parameters of the co-payment allocation problem (CAP). This completes the proof sketch.

One important additional question is whether the bound in Theorem 2 can be improved. We restate here (with modified wording) a proposition from Levi et al. (2017) that answers this question in the negative, showing that the bound is asymptotically tight.

	Optimal Co-payments	**Uniform Co-payments**
q_1	$1-\frac{B(n,c)}{c}$	1
$q_i,\ i\in\{2,\dots,n\}$	$\frac{1}{4}\sqrt{\frac{1}{(n-1)}}-\frac{2B(n,c)}{c}$	0
Q	$1+\frac{n+1}{n-1}-\frac{1}{2}\left(\left(\frac{n+1}{n-1}\right)\sqrt{4+\frac{n-1}{4}}-\sqrt{\frac{n-1}{4}}\right)$	1
y_1	0	$B(n,c)$
$y_i,\ i\in\{2,\dots,n\}$	$\frac{B(n,c)}{(n-1)q_n^C}$	$B(n,c)$

Table 3 Market Equilibria Induced by Optimal and Uniform Co-payments in the Worst-case Instance

PROPOSITION 5 **(Levi et al. (2017))**. *The bound in Theorem 2 is asymptotically tight. Namely, there exists a family of instances of problem (CAP) such that if we let Q^{C} and Q^{U} be the market consumption induced by optimal and uniform co-payments, respectively, then $\frac{Q^{\mathrm{U}}}{Q^{\mathrm{C}}}\xrightarrow{n\to\infty}\frac{1}{2}$.*

The proof of Proposition 5 can be found in the online companion to Levi et al. (2017). It exhibits a family of instances parametrized by the number of firms in the market n. The market equilibria induced by the optimal co-payments and uniform co-payments in these instances are summarized in Table 3, where $B(n,c)=\frac{c}{(n-1)}\left(\sqrt{4+\frac{n-1}{4}}-2\right)$. Similarly to the previous section, this instance consists of one efficient firm and $(n-1)$ homogeneous inefficient firms, facing a constant inverse demand function. Moreover, note from Table 3 that the budget is carefully balanced such that it is optimal to subsidize the inefficient firms only, while giving no co-payment to the efficient firm (see the optimal co-payment structure in Figure 2 and Proposition 2). On the other hand, in the uniform co-payments solution all the subsidies are paid to the efficient firm only, since the inefficient firms are on the verge of start producing and have no output. Namely, the uniform co-payments policy gets the optimal subsidy structure exactly wrong in this instance. Note from Table 3 that the relative performance of uniform co-payments is arbitrarily close to one half as the number of firms n increases. The structure of these instances is very particular, and it is hardly ever obtained from sampling the parameters at random, which partially explains the relatively high worst-case results observed in the simulations presented in Levi et al. (2017).

5.2. A Better Bound for an Important Special Case

In this section, we consider the special case of Cournot competition with linear demand and constant marginal costs. As discussed in Section 2, this is an important model frequently used by researchers as a building block to get detailed insights into complex phenomena. Moreover, the market output in this market equilibrium is equivalent to the one attained by price-taking firms with marginal cost functions $h_i(q_i) = c_i + bq_i$, *i.e.* $g_i = b$ for each firm $i \in \{1, \ldots, n\}$, facing a linear inverse demand function $P(Q) = a - bQ$, see Levi et al. (2017) for the details.

We start by formally stating the main result in this section.

Theorem 3. *For any number $n \geq 2$ of Cournot competitors with constant marginal costs $c_i \geq 0$, for each firm $i \in \{1, \ldots, n\}$, facing a linear decreasing inverse demand function $P(Q) = a - bQ$, where $\sum_{i=1}^{n} q_i = Q$, $a \geq 0$ and $b > 0$, and for any budget $B \geq 0$, let Q^{C} be the market consumption induced by optimal co-payments, and Q^{U} be the market consumption induced by uniform co-payments. Then*

$$\frac{Q^{\mathrm{U}}}{Q^{\mathrm{C}}} \geq \frac{2 + \sqrt{2 + 2/n}}{4} \geq \frac{2 + \sqrt{2}}{4} \approx 85.36\%. \tag{29}$$

Moreover, the first bound in (29) is tight, while the second bound in (29) is asymptotically tight as the number of firms n grows without limit.

Theorem 3 provides a *surprisingly high* worst-case performance guarantee for uniform co-payments in maximizing the market consumption of a good in this model, parametrized by the number of firms in the market n. Moreover, it additionally shows that this guarantee is *tight* and cannot be improved. In particular, if the number of firms is small, namely two or three, then uniform co-payments are guaranteed to induce more than 90% of the optimal market consumption. Furthermore, this leads to a *uniform* bound, showing that the efficiency loss in maximizing market consumption induced by implementing the simple uniform co-payments policy is at most 14.64% *for any instance* of Cournot competition with linear demand and constant marginal costs.

This result is an example that further analysis, based on the structure of important special cases of the general model from Section 3, can lead to a significant improvement in the worst-case performance guarantees for uniform co-payments.

The proof of Theorem 3 follows a more detailed analysis than Theorem 2, in the sense that it provides tight insights into the worst-case performance of uniform co-payments for any finite number of firms. However, the proof is also more involved. For this reason we restrict ourselves to a very simple proof sketch here, and present the details in the e-companion to this paper.

Proof Sketch. First, we derive equations that characterize Q^{C} and Q^{U} in this model, these are analogous to (15) and (16) but specialized to the case of Cournot competition with constant marginal costs and linear demand. We use these to define a mathematical program that minimizes the ratio $\frac{Q^{\mathrm{U}}}{Q^{\mathrm{C}}}$, over the parameters that define a valid instance of the co-payment allocation problem (CAP) for this model. Namely, over the marginal costs c_i for each firm $i \in \{1, \ldots, n\}$, the budget available to the central planner B, and the linear inverse demand function parameters a and b. Then, even if the resulting problem is non-convex, we are able to show that its minimum is equal to $\frac{2+\sqrt{2+2/n}}{4}$. Therefore, the tight bound from Theorem 3 holds. This completes the proof sketch.

6. Conclusions

We study the problem faced by a central planner charging taxes and allocating co-payment subsidies to heterogeneous competing producers of a commodity, with the goal of maximizing its aggregate market consumption subject to a budget constraint. The policy that is most frequently implemented in practical applications of this problem only allocates subsidies and it is uniform, in the sense that every firm gets the same co-payment, even if some firms may be significantly more efficient than others. The central question in this paper is to evaluate the worst possible efficiency loss induced by uniform co-payments, when compared to the optimal firm-dependent co-payments and taxes allocation, as well as to the best policy that allocates only co-payments.

We present the first *worst-case guarantees* for the performance of uniform co-payments in such a model, which provides theoretical support to continue the use of this simple policy in practice. Specifically, we show that uniform co-payments are guaranteed to induce a large fraction of the market consumption induced by the optimal policy that charges taxes and allocates co-payments, in markets with a moderate number of firms. Moreover, uniform co-payments are guaranteed to induce *at least half* of the market consumption induced by optimal co-payments in a fairly general setting, while for an important special case we show an improved guarantee of more than 85% of the market consumption induced by optimal co-payments. In summary, our results suggest that decision makers facing these type of co-payments and taxes allocation problems should not spend time and resources looking into more sophisticated allocation policies, since a very simple uniform co-payment is likely to provide most of their benefits. Future research on this topic should study whether uniform co-payments have a guaranteed performance for a larger family of instances, including additional market competition models.

Appendix

A. Experimental Setup of Section 3

In this Appendix, we provide the details of the experimental setup used to generate the boxplots in Figure 4 and Table 1. We consider the special case of Cournot competition with linear demand and constant marginal costs. Recall, from Section 5.2, that the market output in this market equilibrium is equivalent to the one attained by price-taking firms with marginal cost functions $h_i(q_i) = c_i + bq_i$, *i.e.* $g_i = b$ for each firm $i \in \{1, \dots, n\}$, facing a linear inverse demand function $P(Q) = a - bQ$. Under these assumptions, both (CTAP) and (CAP) are convex optimization problems. Additionally, the market equilibrium induced by uniform co-payments in this model can be computed as the solution to the following convex optimization problem, see Levi et al. (2017).

$$\begin{aligned} & \min_{\mathbf{q},Q} && \sum_{i=1}^{n} c_i q_i + b\frac{q_i^2}{2} - (a - bQ)Q - B\ln(Q) \\ & \text{s.}t. && \sum_{j=1}^{n} q_j = Q \\ (UCAP) \quad & && q_i \geq 0, \ \text{for each } i. \end{aligned}$$

We considered three cases for the number of firms participating in the market, $n \in \{2, 5, 10\}$. For each of them, we generated 100 random instances sampled from the following distributions: the demand parameters a, b, were independent and uniformly distributed in $[0, 50]$, while the marginal costs c_i were independent and uniformly distributed in $[0, a]$, for each firm i. In order to incorporate uncertainty into the central planner's marginal costs estimates, we considered two cases for the uncertainty level $\gamma \in \{0.1, 0.2\}$. For each of the 100 true marginal costs c_i, we generated 30 central planner's estimates $\tilde{c}_i$, sampled at random from the support $[c_i(1-\gamma), c_i(1+\gamma)]$, for each $\gamma \in \{0.1, 0.2\}$. For each of the 3,000 central planner's estimates $\tilde{\mathbf{c}}$, we computed the optimal co-payments and taxes $\mathbf{y}^{CT}(\tilde{\mathbf{c}})$, optimal co-payments $\mathbf{y}^{C}(\tilde{\mathbf{c}})$, and uniform co-payments $\mathbf{y}^{U}(\tilde{\mathbf{c}})$, by solving problems (CTAP), (CAP), and (UCAP), respectively, *based on the marginal cost estimates*. Finally, we used the associated true marginal cost $\mathbf{c}$ to compute the true induced market consumptions $Q^{CT}(\gamma)$, $Q^{C}(\gamma)$, and $Q^{U}(\gamma)$. In order to solve all the convex programs in these simulations we used CVX, see Grant and Boyd (2012).

Acknowledgments

The research of the first author was partially supported by National Science Foundation Grants CMMI-0846554 (CAREER Award) and CMMI-1537536 and an AFOSR award FA9550-11-1-0150. The research of the second author was partially supported by the MIT Energy Initiative Seed funding as well as NSF grant CMMI-1162034. The research of the third author was partially supported by the Connaught New Researcher Award, University of Toronto. The authors thank Amir Beck (the department editor handling this paper) as well as the anonymous associate editor and reviewers for their comments, which helped to substantially improve both the content and exposition of this work.

References

A2S2. 2015. Who/ema prequalified act manufacturers. URL `http://www.a2s2.org/market-data/map-of-stakeholders/\#1`.

AMFm Independent Evaluation Team. 2012. Independent evaluation of phase 1 of the affordable medicines facility - malaria (amfm), multi-country independent evaluation report. Tech. rep., The Global Fund.

Arrow, K. J., C. B. Panosian, H. Gelband. 2004. *Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance*. National Academies Press, Washington, DC.

Aydin, G., E. L. Porteus. 2009. Manufacturer-to-retailer versus manufacturer-to-consumer rebates in a supply chain. *Retail Supply Chain Management*. Springer, 237–270.

Bertsekas, D. P. 1999. *Nonlinear programming*. Athena Scientific, Belmont, MA.

Bertsimas, D., J. N. Tsitsiklis. 1997. *Introduction to linear optimization*. Athena Scientific Belmont, MA.

Boulton, I. 2011. Background paper: Forum for africa-based pharmaceutical manufacturers and potential investors. Tech. rep., The Global Fund. URL `http://www.theglobalfund.org/documents/amfm/AMFm_ForumForAfricaBasedPharmaceuticalManufacturersAndPotentialInvestors_Background_Paper_en/`.

Cachon, G. P. 2003. Supply chain coordination with contracts. A. G. de Kok, S. C. Graves, eds., *Handbooks in Operations Research and Management Science, Vol 11: Supply Chain Mgt*. Elsevier, 229 – 339.

Charnes, A., W. W. Cooper. 1962. Programming with linear fractional functionals. *Naval Research Logistics Quarterly* **9**(3-4) 181196.

Cohen, M., R. Lobel, G. Perakis. 2016. The impact of demand uncertainty on consumer subsidies for green technology adoption. *Management Science* **62**(5) 1235–1258.

Corbett, C. J., U. S. Karmarkar. 2001. Competition and structure in serial supply chains with deterministic demand. *Management Science* **47**(7) 966–978.

Deo, S., C. J. Corbett. 2009. Cournot competition under yield uncertainty: The case of the u.s. influenza vaccine market. *Manufacturing and Service Operations Management* **11**(4) 563 – 576.

Farahat, A., G. Perakis. 2011. Technical note-a comparison of bertrand and cournot profits in oligopolies with differentiated products. *Operations research* **59**(2) 507–513.

Fullerton, D., G. E. Metcalf. 2002. Chapter 26 tax incidence. Alan J. Auerbach, Martin Feldstein, eds., *Handbook of Public Economics*, vol. 4. Elsevier, 1787 – 1872.

Gallien, Jeremie, Iva Rashkova, Rifat Atun, Prashant Yadav. 2017. National drug stockout risks and the global fund disbursement process for procurement. *Production and Operations Management* **26**(6) 997–1014.

Grant, M., S. Boyd. 2012. CVX: Matlab software for disciplined convex programming, version 2.0 beta. `http://cvxr.com/cvx`.

Krass, D., T. Nedorezov, A. Ovchinnikov. 2013. Environmental taxes and the choice of green technology. *Production and Operations Management* **22**(5) 1035–1055.

Levi, R., G. Perakis, G. Romero. 2017. On the effectiveness of uniform subsidies in increasing market consumption. *Management Science* **63**(1) 40–57.

Marcotte, P., M. Patriksson. 2006. Traffic equilibrium. C. Barnhart, G. Laporte, eds., *Handbooks in Operations Research and Management Science, Vol. 14: Transportation*. Elsevier, 623 – 713.

Raz, Gal, Anton Ovchinnikov. 2015. Coordinating pricing and supply of public interest goods using government rebates and subsidies. *IEEE Transactions on Engineering Management* **62**(1) 65–79.

Taylor, T. A., W. Xiao. 2014. Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. *Management Science,* **60**(10) 2461–2477.

Tirole, Jean. 1988. *The theory of industrial organization*. MIT press.

Vives, X. 2001. *Oligopoly Pricing: Old Ideas and New Tools*. The MIT press, Cambridge, MA.

E-Companion: Proofs

EC.1. Proofs of Section 4

PROPOSITION EC.1. *Let $t, u \in \{1, \dots, n\}$ be indexes as defined in Propositions 1 and 3, respectively. Then, for any number of firms $n \geq 2$, the worst-case performance of uniform co-payments must be attained at an instance of the problem (CTAP) such that $t = u = n$.*

The proof of proposition EC.1 is analogous to the first part of proposition EC.2, and it is therefore omitted.

LEMMA EC.1. *For any $n \geq 2$, n-dimensional vectors $\mathbf{c}$, $\boldsymbol{g}$ such that $0 \leq c_1 \leq \dots \leq c_n$, $0 < g_1 \leq \dots \leq g_n$, it must be the case that,* $(n-1)\left(\sum_{i=1}^n \frac{c_n - c_i}{g_i}\right)^2 + \left(\sum_{i=1}^n \frac{c_i}{g_i}\right)^2 - \sum_{i=1}^n \frac{c_i^2}{g_i} \sum_{i=1}^n \frac{1}{g_i} \geq 0.$

Proof. For any $n \geq 2$, and fixed $0 < g_1 \leq \dots \leq g_n$, the left hand side of the inequality in the lemma is a continuous function of $\mathbf{c}$. Moreover, for fixed c_i, $i \in \{1, \dots, n-1\}$ it grows without bound as c_n goes to infinity. It follows that it must attain a minimum in the feasible set defined by $0 \leq c_1 \leq \dots \leq c_n$.

We show that any minimizer of the left hand side of the inequality in the lemma, over $0 \leq c_1 \leq \dots \leq c_n$, must be such that each c_i is either equal to 0 or to c_n. First, assume for a contradiction that there exists a minimizer such that at least two indexes have values strictly between 0 and c_n. Let i be the first index such that $c_i > 0$, and j be the last index such that $c_j < c_n$, then it follows that we can increase c_j and decrease c_i by the same $\epsilon > 0$ small enough, and obtain a strictly smaller value of the left hand side of the inequality in the lemma, a contradiction. Hence, there can be at most one index $i \in \{1 \dots, n\}$ such that $0 < c_i < c_n$. However, note that then the derivative of the left hand side of the inequality in the lemma, with respect to c_i is $\frac{2}{g_i}\left(c_n\left(\sum_{j=1}^n \frac{1}{g_j} - \sum_{j=1}^i \frac{n}{g_j}\right) + c_i\left(\frac{n}{g_i} - \sum_{j=1}^n \frac{1}{g_j}\right)\right) < 0$. Where the inequality follows from $0 < g_1 \leq \dots \leq g_n$ and $c_i < c_n$. Therefore, we can increase c_i and obtain a strictly smaller value of the left hand side of the inequality in the lemma, a contradiction. Hence, any minimum must satisfy that there exists an index $k \in \{0, \dots, n-1\}$ such that $c_i = 0$ for $i \in \{1, \dots, k\}$, and $c_i = c_n$ for $i \in \{k+1, \dots, n\}$.

To conclude, note that the inequality in the lemma is satisfied in this case. Specifically, it evaluates to $(n-1)\left(\sum_{i=1}^{k}\frac{c_n}{g_i}\right)^2+\left(\sum_{i=k+1}^{n}\frac{c_n}{g_i}\right)^2-\sum_{i=k+1}^{n}\frac{c_n^2}{g_i}\sum_{i=1}^{n}\frac{1}{g_i}\geq 0$, whose left hand side is increasing in the index k, and the inequality is tight for the case $k=0$, for any $c_n\geq 0$ and $0<g_1\leq\ldots\leq g_n$. $\square$

Proof of Proposition 4. For any $n\geq 2$, consider any parameter $c>0$ and the instance of problem (CTAP) defined by the central planner having no budget, i.e. $B=0$, for price-taking firms with affine increasing marginal costs $h_1(q_1)=cq_1$, $h_i(q_i)=c+cq_i$ for each $i\in\{2,\ldots,n\}$, facing a constant inverse demand function, $P=c$.

We now check that the unique market equilibrium induced by the optimal co-payments and taxes, and by optimal co-payments, in this instance are the ones in Table 2 in the paper.

From $B=0$ it follows that the market equilibrium induced by any policy that only allocates co-payments is the solution to a convex program in this setting, and thus unique, see Levi et al. (2017) for the details. Therefore, it is enough to check that the market equilibrium in the third column of Table 2 satisfies equations (16)-(18) in Proposition 3, which is trivially true.

On the other hand, problem (CTAP) is also a convex program in this case, therefore the market equilibrium induced by optimal co-payments and taxes is unique and the KKT conditions from Proposition 1, together with the feasibility constraints, are necessary and sufficient for optimality. We first check feasibility. The budget is exhausted in constraint (2) since $\sum_{i=1}^{n}q_i^{\mathrm{CT}}y_i^{\mathrm{CT}}=\frac{-(n-1)cq_n^{\mathrm{CT}}}{1+\sqrt{n}}q_1^{\mathrm{CT}}+(n-1)\frac{cq_1^{\mathrm{CT}}}{1+\sqrt{n}}q_n^{\mathrm{CT}}=0=B$, and the market equilibrium condition (5) is also satisfied. Specifically, for firm 1 we have $h_1(q_1^{\mathrm{CT}})-y_1^{\mathrm{CT}}=c\left(q_1^{\mathrm{CT}}+\frac{-(n-1)q_n^{\mathrm{CT}}}{1+\sqrt{n}}\right)=\left(\frac{c}{1+\sqrt{n}}\right)(\sqrt{n}q_1^{\mathrm{CT}}+Q^{\mathrm{CT}})=\left(\frac{c}{1+\sqrt{n}}\right)2Q^{\mathrm{CT}}=c=P$, while for firm $i\in\{2,\ldots,n\}$ we have $P=cq_1^{\mathrm{CT}}-y_1^{\mathrm{CT}}=\left(\frac{c}{2}+cq_i^{\mathrm{CT}}\right)-\left(y_i^{\mathrm{CT}}-\frac{c}{2}\right)=c+cq_i^{\mathrm{CT}}-y_i^{\mathrm{CT}}=h_i(q_i^{\mathrm{CT}})-y_i^{\mathrm{CT}}$, where the first equality follows as before, and the second equality follows from the values of q_1^{CT}, q_n^{CT}, y_1^{CT} and y_n^{CT}. In a similar way, equations (6)-(9) in Proposition 1 can be verified with some algebra. This completes the proof of the first statement in the proposition.

For the second statement in the proposition, for any any parameter $c>0$ let $\epsilon\in(0,c)$, and consider the following instance with $n=2$ firms. Assume that the central planner has no budget, i.e. $B=0$, and the price-taking firms have marginal costs $h_1(q_1)=cq_1$, $h_2(q_2)=c+\epsilon$ (although $h_2(q_2)$ is constant, it can be made strictly increasing with negligible slope without changing the result), and face a constant inverse demand function, $P=c$.

It is not hard to see that the unique market equilibrium induced by any policy that allocates co-payments only are the ones in the third column of Table 2. In particular, $Q^C=1$.

Similarly, the unique market equilibrium induced by optimal co-payments and taxes is such that $q_1^{CT}=\frac{c+\epsilon}{2c}$, $h_1(q_1^{CT})=\frac{c+\epsilon}{2}$, $y_1^{CT}=-\frac{c-\epsilon}{2}$, and $q_2^{CT}=\left(\frac{c+\epsilon}{2c}\right)\left(\frac{c-\epsilon}{2\epsilon}\right)$, $h_2(q_2^{CT})=c+\epsilon$, $y_2^{CT}=\epsilon$.

Hence, $\frac{Q^{CT}}{Q^C}=Q^{CT}=\left(\frac{c+\epsilon}{2c}\right)\left(\frac{c+\epsilon}{2\epsilon}\right)\xrightarrow{\epsilon\to 0}\infty$. This completes the proof. □

EC.2. Proofs of Section 5.1

We formalize the proof sketch of Theorem 2, it follows four steps. First, Proposition EC.2 re-writes constraints (22)-(24) in problem (P^n_{lmu}) in terms of the variables of the problem $(\mathbf{c},\mathbf{g},B,P())$, while also significantly reducing the number of indexes l,m,u,n that we need to consider. This allows to re-write problem (P^n_{lmu}), which is in implicit form, as problem (P_l^G) below. Moreover, Proposition EC.2 asserts that it is sufficient to show that problem (P_l^G) has a non-negative optimal objective value for each $l\in\{1,\dots,n\}$, $n\geq 2$, for the bound of one half to hold. Then, Proposition EC.3 shows that any optimal solution of problem (P_l^G) must satisfy three properties. Namely, (i) it must have a constant inverse demand function, (ii) it must be such that $q_n^{\mathrm{U}}=0$, and (iii) it must be such that either $y_l^{\mathrm{C}}=0$ or $y_l^{\mathrm{C}}=\frac{c_l-c_{l-1}}{2}$. Finally, the proof of Theorem 2 shows that any feasible solution of problem (P_l^G) that satisfies these three properties must have a non-negative objective value.

Proposition EC.2. *Let $l,m,u\in\{1,\dots,n\}$ be indexes as defined in Propositions 2 and 3, respectively. Then, for any number of firms $n\geq 2$, the worst-case performance of uniform co-payments must be attained in an instance of problem (CAP) such that $l\leq m=u=n$. Moreover, these instances are characterized by the following constraints*

$$y_l^{\mathrm{C}}\geq 0 \Leftrightarrow B-\sum_{i=l}^{n}\left(\frac{2P(Q^{\mathrm{C}})-c_i-c_l}{2}\right)\left(\frac{c_i-c_l}{2g_i}\right)\geq 0. \tag{EC.1}$$

$$y_l^{\mathrm{C}} \leq \frac{c_l - c_{l-1}}{2} \Leftrightarrow \sum_{i=l}^{n} \left(\frac{2P(Q^{\mathrm{C}}) - c_i - c_{l-1}}{2} \right) \left(\frac{c_i - c_{l-1}}{2g_i} \right) - B \geq 0. \tag{EC.2}$$

$$q_n^{\mathrm{C}} \geq 0 \Leftrightarrow B - \sum_{i=l}^{n} \left(\frac{c_n - c_i}{2} \right) \left(\frac{c_n + c_i - 2P(Q^{\mathrm{C}})}{2g_i} \right) \geq 0. \tag{EC.3}$$

$$q_n^{\mathrm{U}} \geq 0 \Leftrightarrow B - \left(c_n - P(Q^{\mathrm{U}})\right) \left(\sum_{i=1}^{n} \frac{c_n - c_i}{g_i} \right) \geq 0. \tag{EC.4}$$

Hence, under the same assumptions of Theorem 2, if problem (P_l^G) *has a non-negative optimal objective value for each* $l \in \{1, \ldots, n\}$*,* $n \geq 2$*, then it follows that* $2Q^{\mathrm{U}} \geq Q^{\mathrm{C}}$.

Proof. For any $n \geq 2$, consider an instance of problem (CAP) and let $l, m, u \in \{1, \ldots, n\}$ be the indexes defined in Propositions 2 and 3, respectively (see Figures 2 and 3).

First, assume that $u < m$. We show that there exists a modified instance of problem (CAP) that attains a strictly worst performance of uniform co-payments. Let $(\mathbf{q}^{\mathrm{U}}, Q^{\mathrm{U}})$ be the solution induced by uniform co-payments in the original instance. Then, at the market equilibrium induced by uniform co-payments we must have $c_u + g_u q_u^{\mathrm{U}} = P(Q^{\mathrm{U}}) + \frac{B}{Q^{\mathrm{U}}} < c_m$. Let $\hat{\imath}$ be the first index such that $\hat{\imath} \geq l$, and $c_u < c_{\hat{\imath}}$. Note that $\hat{\imath} \in \{l, \ldots, m\}$, then $c_{\hat{\imath}} > P(Q^{\mathrm{U}}) + \frac{B}{Q^{\mathrm{U}}}$. It follows that we can reduce the value of $c_{\hat{\imath}}$, by $\epsilon > 0$ sufficiently small, without affecting the uniform co-payments solution $(\mathbf{q}^{\mathrm{U}}, Q^{\mathrm{U}})$, while obtaining a strictly larger value for the aggregate consumption induced by optimal co-payments. Specifically, let $(\mathbf{q}^{\mathrm{C}}, Q^{\mathrm{C}})$ be an optimal solution to problem (CAP) for the original instance where $u < m$. Note that $(\mathbf{q}^{\mathrm{C}}, Q^{\mathrm{C}})$ is feasible for the modified instance where we reduce the value of $c_{\hat{\imath}}$, by $\epsilon > 0$ sufficiently small. Moreover, from the budget constraint not being binding for $(\mathbf{q}^{\mathrm{C}}, Q^{\mathrm{C}})$ in the modified instance, it follows that we can increase some q_i^{C}, and Q^{C}, by $\delta > 0$ small enough, maintain the feasibility for problem (CAP), and obtain a strictly larger objective value, reducing the relative performance of uniform co-payments.

Now assume that $u > m$. We again show that there exists a modified instance of problem (CAP) that attains a strictly smaller value of the ratio $Q^{\mathrm{U}}/Q^{\mathrm{C}}$. Note that $u > m$ implies $c_u > c_m$ and $q_u^{\mathrm{U}} > 0$. Therefore, we can increase the value of c_u by $\epsilon > 0$ sufficiently small, without changing the optimal co-payments solution $(\mathbf{q}^{\mathrm{C}}, Q^{\mathrm{C}})$, while decreasing Q^{U}. Specifically, from $u = n$ and $q_u^{\mathrm{U}} > 0$

it follows that the set of active firms at the market equilibrium does not change. Additionally, the derivatives of the right hand side of equation (16) with respect to c_u and $Q^{\rm U}$, respectively, are

$$-\frac{1}{2g_u}-\frac{\frac{1}{2g_u}\sum_{i=1}^{u}\frac{P(Q^{\rm U})-c_i}{2g_i}}{\sqrt{\left(\sum_{i=1}^{u}\frac{P(Q^{\rm U})-c_i}{2g_i}\right)^2+\sum_{i=1}^{u}\frac{B}{g_i}}}<0,\ \text{and}\ \sum_{i=1}^{u}\frac{P'(Q^{\rm U})}{2g_i}+\frac{\sum_{i=1}^{u}\frac{P'(Q^{\rm U})}{2g_i}\sum_{i=1}^{u}\frac{P(Q^{\rm U})-c_i}{2g_i}}{\sqrt{\left(\sum_{i=1}^{u}\frac{P(Q^{\rm U})-c_i}{2g_i}\right)^2+\sum_{i=1}^{u}\frac{B}{g_i}}}<0.$$

Hence, $Q^{\rm U}$ decreases, strictly reducing the relative performance of uniform co-payments with respect to the original instance.

The remainder of the proof is omitted for the sake of brevity. It consists on verifying that constraints (EC.1) and (EC.2) are equivalent to the non-negativity of the co-payments (10), and constraints (EC.3) and (EC.4) are equivalent to the non-negativity of the market outputs (4).

To conclude, note that the equality in the objective function of problem (P_l^G) follows from the fixed point equations (15) for $Q^{\rm C}$ and (16) for $Q^{\rm U}$, for the case $u=m=n$. □

From Proposition EC.2 it follows that there is no loss of generality in focusing on problem (P_{lnn}^n) (i.e., $m=u=n$) for each index $l\in\{1,\dots,n\}$. Recall that l is the index of the first firm that receives a positive co-payment in the optimal co-payment allocation, see Figure 2. Moreover, we can re-write problem (P_{lnn}^n) as problem (P_l^G) below, which only depends on $P()$ through $P(Q^{\rm U})$ and $P(Q^{\rm C})$.

$$\begin{aligned}
\min_{\mathbf{c},\mathbf{g},B,P(Q^{\rm U}),P(Q^{\rm C})}\quad & 2Q^{\rm U}-Q_l^{\rm C}=2\sqrt{\left(\sum_{i=1}^{n}\frac{P(Q^{\rm U})-c_i}{2g_i}\right)^2+\sum_{i=1}^{n}\frac{B}{g_i}}+\sum_{i=l}^{n}\frac{P(Q^{\rm C})-c_i}{2g_i}+\sum_{i=1}^{n}\frac{P(Q^{\rm U})-P(Q^{\rm C})}{g_i}\\
&-\sqrt{\left(\sum_{i=l}^{n}\frac{P(Q^{\rm C})-c_i}{2g_i}\right)^2+\sum_{i=l}^{n}\frac{B}{g_i}+\sum_{i=l}^{n}\frac{1}{g_i}\left(\sum_{i=l}^{n}\left(\frac{c_i-c_l}{2}\right)^2\frac{1}{g_i}\right)-\left(\sum_{i=l}^{n}\frac{c_i-c_l}{2g_i}\right)^2}\\
\text{s.}t.\quad & \text{constraints } (25)-(27) \text{ and } (EC.1)-(EC.4) \text{ are satisfied}\\
(\mathrm{P}_l^G)\quad & P(Q^{\rm U})\geq P(Q^{\rm C})\geq 0. \qquad\qquad \text{(EC.5)}
\end{aligned}$$

Proposition EC.3 below identifies conditions that any minimizer of problem (P_l^G) must satisfy.

Proposition EC.3. *For any $n\geq 2$, and for each $l\in\{1,\dots,n\}$, any minimizer of (P_l^G) must satisfy three properties: (i) it must satisfy $P(Q^{\rm U})=P(Q^{\rm C})=P$, (ii) it must be such that constraint (EC.4) is tight, and (iii) it must be such that either constraint (EC.1) or constraint (EC.2) is also* *tight. Moreover, when focusing on instances such that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ constraint (EC.3) is redundant.*

The proof of Proposition EC.3 requires the analysis of several sub-cases. For the sake of clarity, we split it into several steps given by Propositions EC.4-EC.7 stated below.

Proposition EC.4. *For any $n \geq 2$, and for each $l \in \{1,\dots,n\}$, any minimizer of the mathematical program (P_l^G) must be such that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$. Moreover, when focusing on these instances constraint (EC.3) becomes redundant.*

Proof. The objective function of problem (P_l^G) is increasing in $P(Q^{\mathrm{U}})$ and decreasing in $P(Q^{\mathrm{C}})$. Additionally, the left hand side of constraint (EC.4) is increasing in $P(Q^{\mathrm{U}})$, while the left hand side of constraint (EC.1) is decreasing in $P(Q^{\mathrm{C}})$, and the left hand side of constraints (EC.2) and (EC.3) are increasing in $P(Q^{\mathrm{C}})$. Hence, in order to study the worst-case performance of uniform co-payments there is no loss of generality in focusing on instances of the problem (CAP) such that either $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$, or both (EC.4) and (EC.1) hold with equality. To conclude, we show that in the latter case there is no loss of generality in focusing on the case where $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ either.

Consider any instance of the problem (CAP) such that both (EC.4) and (EC.1) hold with equality. Namely, such that $q_n^{\mathrm{U}} = 0$ and $y_l^{\mathrm{C}} = 0$ (see Proposition EC.2). From equation (14) for $m = n$ it follows that $y_l^{\mathrm{C}} = 0$ implies

$$\sqrt{\left(\sum_{i=l}^{n} \frac{P(Q^{\mathrm{C}}) - c_i}{2g_i}\right)^2 + \sum_{i=l}^{n} \frac{B}{g_i} + \sum_{i=l}^{n} \frac{1}{g_i} \left(\sum_{i=l}^{n} \left(\frac{c_i - c_l}{2}\right)^2 \frac{1}{g_i}\right) - \left(\sum_{i=l}^{n} \frac{c_i - c_l}{2g_i}\right)^2} = \sum_{i=l}^{n} \frac{P(Q^{\mathrm{C}}) - c_l}{2g_i}. \tag{EC.6}$$

Similarly, from (16) and (17) for $u = n$ it follows that $q_n^{\mathrm{U}} = 0$ implies

$$\sqrt{\left(\sum_{i=1}^{n} \frac{P(Q^{\mathrm{U}}) - c_i}{2g_i}\right)^2 + \sum_{i=1}^{n} \frac{B}{g_i}} = \sum_{i=1}^{n} \frac{2c_n - P(Q^{\mathrm{U}}) - c_i}{2g_i}. \tag{EC.7}$$

From (EC.6) and (EC.7) it follows that if both (EC.4) and (EC.1) hold with equality, then the objective function of problem (P_l^G) simplifies to $\sum_{i=1}^{n} \frac{2c_n - P(Q^{\mathrm{C}}) - c_i}{g_i} - \sum_{i=l}^{n} \frac{c_i - c_l}{2g_i}$, which is decreasing in $P(Q^{\mathrm{C}})$ and independent of $P(Q^{\mathrm{U}})$. Moreover, that both (EC.4) and (EC.1) hold with equality implies $(c_n - P(Q^{\mathrm{U}}))\left(\sum_{i=1}^{n}\frac{c_n-c_i}{g_i}\right) = B = \sum_{i=l}^{n}\left(\frac{2P(Q^{\mathrm{C}})-c_i-c_l}{2}\right)\left(\frac{c_i-c_l}{2g_i}\right) \geq 0$. Namely, in order the worst-case performance of uniform co-payments, if both (EC.4) and (EC.1) hold with equality and $P(Q^{\mathrm{U}}) > P(Q^{\mathrm{C}})$, then we can always increase $P(Q^{\mathrm{C}})$ and decrease $P(Q^{\mathrm{U}})$ appropriately until $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$.

To conclude note that from $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ it follows that constraint (EC.3) is redundant. Specifically, if $c_n < P$ then $B \geq 0 > \sum_{i=l}^{n}\left(\frac{c_n-c_i}{2}\right)\left(\frac{c_n+c_i-2P}{2g_i}\right)$, that is (25) implies (EC.3). Similarly, if $c_n \geq P$ then $B \geq (c_n - P)\left(\sum_{i=1}^{n}\frac{c_n-c_i}{g_i}\right) \geq \sum_{i=l}^{n}\left(\frac{c_n-c_i}{2}\right)\left(\frac{c_n+c_i-2P}{2g_i}\right)$, and (EC.4) implies (EC.3). This completes the proof. □

PROPOSITION EC.5. *For any $n \geq 2$, and for each $l \in \{1,\ldots,n\}$, any minimizer of the mathematical program* (P_l^G) *must be such that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$, and at least one of the constraints (EC.1), (EC.2), or (EC.4), is tight.*

Proof. For any $n \geq 2$, $l \in \{1,\ldots,n\}$, from Proposition EC.4 it follows that we can focus, without loss of generality, on feasible solutions of (P_l^G) such that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$. Consider any n-dimensional vectors $\mathbf{c}$ and $\mathbf{g}$ where $0 \leq c_1 \leq \ldots \leq c_n$, and $g_i > 0$ for each $i \in \{1,\ldots,n\}$, such that the set of values of (B,P) that satisfy the constraints (25) and (EC.1)-(EC.4) in problem (P_l^G) is non-empty, and denote by $f(B,P)$ its objective function. Note that $f(B,P)$ is continuous and coercive. Additionally, the constraints (25) and (EC.1)-(EC.4) define a closed feasible set for (B,P), ensuring the existence of a minimum of the function $f(B,P)$, see for example Bertsekas (1999).

Moreover, note that a minimum cannot be attained at an interior point of the feasible set. Specifically, we have that $\frac{\partial f(B,P)}{\partial B} = 0$ implies $\frac{\partial f(B,P)}{\partial P} > 0$. Therefore, the minimum must be attained at a point (B,P) such that at least one of the constraints (25) and (EC.1)-(EC.4) is binding.

However, from Proposition EC.4 we also know that constraint (EC.3) is redundant. To conclude, note that that if constraint (25) is tight, then $B = 0$ and $f(0,P)$=1, namely $Q^{\mathrm{U}} = Q^{\mathrm{C}}$. This is actually a maximizer of $f(B,P)$. Hence, at least one of the constraints (EC.1), (EC.2), or (EC.4) must be binding at any minimizer of (P_l^G). □

Proposition EC.6. *For any $n \geq 2$, and for each $l \in \{1, \ldots, n\}$, any feasible solution of the mathematical program (P_l^G) such that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ and constraint (EC.1), or constraint (EC.2), is tight, is dominated by a feasible solution where additionally constraint (EC.4) is also tight.*

Proof. For any $n \geq 2$, $l \in \{1, \ldots, n\}$, assume $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ and consider any n-dimensional vectors $\mathbf{c}$ and $\mathbf{g}$ where $0 \leq c_1 \leq \ldots \leq c_n$, and $g_i > 0$ for each $i \in \{1, \ldots, n\}$, such that the set of values of P that satisfy constraint (EC.1) with equality, and constraints (25) and (EC.2)-(EC.4) in problem (P_l^G) is non-empty. We show that any minimum over P of the objective function of problem (P_l^G) in this set must also satisfy constraint (EC.4) with equality.

Note that from constraint (EC.1) being tight we can rewrite constraint (EC.4) as

$$P \geq \frac{c_n \sum_{i=1}^n \frac{c_n - c_i}{g_i} + \sum_{i=l}^n \left(\frac{c_i + c_l}{2}\right)\left(\frac{c_i - c_l}{2gi}\right)}{\sum_{i=1}^n \frac{c_n - c_i}{g_i} + \sum_{i=l}^n \frac{c_i - c_l}{2gi}}, \tag{EC.8}$$

which implies $P \geq c_l \geq 0$, as well as (EC.2) and (25). Moreover, from Proposition EC.4 it follows that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ implies that constraint (EC.3) is redundant. Therefore, the feasibility of the instance is defined by the inequality (EC.8) only.

On the other hand, constraint (EC.1) being tight is equivalent to $y_l^{\mathrm{C}} = 0$ (see Proposition EC.2). From equation (14) for $m = n$, together with the expression for B from (EC.1) being tight, it follows that the objective function of problem (P_l^G) simplifies to $2\sqrt{\left(\sum_{i=1}^n \frac{P - c_i}{2g_i}\right)^2 + \sum_{i=1}^n \frac{1}{g_i} \sum_{i=l}^n \left(\frac{2P - c_i - c_l}{2g_i}\right)\left(\frac{c_i - c_l}{2}\right)} - \sum_{i=l}^n \frac{c_i - c_l}{2g_i}$. Note that this expression is continuous and increasing in P. Therefore, it attains its minimum at the lower bound of the interval defined by inequality (EC.8), or equivalently, when (EC.4) also holds with equality.

The proof for the case where constraint (EC.2) is tight instead is analogous, and it is therefore omitted for the sake of brevity. □

Proposition EC.7. *For any $n \geq 2$, and for each $l \in \{1, \ldots, n\}$, any feasible solution of the mathematical program (P_l^G) such that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ and constraint (EC.4) is tight, is dominated by a feasible solution where additionally one of the constraints (EC.1) or (EC.2) is also tight.*

Proof. For any $n \geq 2$, $l \in \{1,\ldots,n\}$, assume $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ and consider any n-dimensional vectors $\mathbf{c}$ and $\mathbf{g}$ where $0 \leq c_1 \leq \ldots \leq c_n$, and $g_i > 0$ for each $i \in \{1,\ldots,n\}$, such that the set of values of P that satisfy constraint (EC.4) with equality, and constraints (25) and (EC.1)-(EC.4) in problem (P_l^G) is non-empty. We show that any minimum over P of the objective function of problem (P_l^G) in this set must be such that one of the constraints (EC.1) or (EC.2) is also tight.

Note that from constraint (EC.4) being tight we can rewrite constraint (EC.1) as

$$P \leq \frac{c_n \sum_{i=1}^n \frac{c_n - c_i}{g_i} + \sum_{i=l}^n \left(\frac{c_i + c_l}{2}\right)\left(\frac{c_i - c_l}{2gi}\right)}{\sum_{i=1}^n \frac{c_n - c_i}{g_i} + \sum_{i=l}^n \frac{c_i - c_l}{2gi}}. \tag{EC.9}$$

Additionally, from (EC.4) being tight, it follows that (25) is equivalent to $P \leq c_n$, which is redundant with (EC.9). Similarly, from (EC.4) being tight we can rewrite constraint (EC.2) as

$$P \geq \frac{c_n \sum_{i=1}^n \frac{c_n - c_i}{g_i} + \sum_{i=l}^n \left(\frac{c_i + c_{l-1}}{2}\right)\left(\frac{c_i - c_{l-1}}{2gi}\right)}{\sum_{i=1}^n \frac{c_n - c_i}{g_i} + \sum_{i=l}^n \frac{c_i - c_{l-1}}{2gi}}, \tag{EC.10}$$

which implies $P \geq 0$. Moreover, from Proposition EC.4 it follows that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ implies that constraint (EC.3) is redundant. Therefore, the feasibility of the instance is defined by the inequalities (EC.9) and (EC.10) only.

On the other hand, constraint (EC.4) being tight is equivalent to $q_n^{\mathrm{U}} = 0$ (see Proposition EC.2). From (16) and (17) for $u = n$, together with the expression for B from (EC.4) being tight, it follows that the objective function of problem (P_l^G) simplifies to

$$\sum_{i=1}^n \frac{2c_n - P - c_i}{g_i} + \sum_{i=l}^n \frac{P - c_i}{2g_i}$$
$$- \sqrt{\left(\sum_{i=l}^n \frac{P - c_i}{2g_i}\right)^2 + \sum_{i=l}^n \frac{c_n - P}{g_i} \sum_{i=1}^n \frac{c_n - c_i}{g_i} + \sum_{i=l}^n \frac{1}{g_i} \sum_{i=l}^n \left(\frac{c_i - c_l}{2}\right)^2 \frac{1}{g_i} - \left(\sum_{i=l}^n \frac{c_i - c_l}{2g_i}\right)^2}, \tag{EC.11}$$

which is concave in P. Moreover, the simplified expressions (EC.9) and (EC.10) define a non-empty compact interval for P. Therefore, it follows that the objective function of problem (P_l^G) must attain its minimum at one of the extremes of the feasible interval. Equivalently, when at least one of (EC.1) or (EC.2) also holds with equality. □

Now we are ready to complete the proof of the main result in Section 5.1.

Proof of Theorem 2. From Propositions EC.2 and EC.3, it follows that it is enough to show that for any $n \geq 2$, $l \in \{1,\dots,n\}$, any feasible solution to the problem (P_l^G), such that (i) $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$, (ii) constraint (EC.4) is tight, and (iii) one of constraints (EC.1) or (EC.2) is also tight, must have a non-negative objective value. Importantly, under these conditions the objective function of (P_l^G) simplifies significantly (see equation (EC.15) below), which allows to show that it is non-negative based on first principles. We show here the first case, where constraint (EC.1) is tight. Note that from constraints (EC.4) and (EC.1) being tight it follows that the value of P is uniquely defined by

$$(c_n - P)\left(\sum_{i=1}^{n} \frac{c_n - c_i}{g_i}\right) = B = \sum_{i=l}^{n} \left(\frac{2P - c_i - c_l}{2}\right)\left(\frac{c_i - c_l}{2g_i}\right) \geq 0, \tag{EC.12}$$

for any $0 \leq c_1 \leq \ldots \leq c_n$ and $g_i > 0$, $i \in \{1,\dots,n\}$. The inequality in (EC.12) follows from (25).

Inequality (EC.12) implies $c_n > P > 0$, as well as constraint (EC.2). Moreover, from Proposition EC.3 it follows that $P(Q^{\mathrm{U}}) = P(Q^{\mathrm{C}}) = P$ implies that constraint (EC.3) is redundant. Therefore, the feasibility of the instance is only determined by (EC.12).

On the other hand, constraint (EC.1) being tight is equivalent to $y_l^{\mathrm{C}} = 0$ (see Proposition EC.2). From equation (14) for $m = n$ it follows that $y_l^{\mathrm{C}} = 0$ implies

$$\sqrt{\left(\sum_{i=l}^{n} \frac{P - c_i}{2g_i}\right)^2 + \sum_{i=l}^{n} \frac{B}{g_i} + \sum_{i=l}^{n} \frac{1}{g_i}\left(\sum_{i=l}^{n}\left(\frac{c_i - c_l}{2}\right)^2 \frac{1}{g_i}\right) - \left(\sum_{i=l}^{n} \frac{c_i - c_l}{2g_i}\right)^2} = \sum_{i=l}^{n} \frac{P - c_l}{2g_i}. \tag{EC.13}$$

Similarly, constraint (EC.4) being tight is equivalent to $q_n^{\mathrm{U}} = 0$ (see Proposition EC.2). From equations (16) and (17) for $u = n$ it follows that $q_n^{\mathrm{U}} = 0$ implies

$$\sqrt{\left(\sum_{i=1}^{n} \frac{P - c_i}{2g_i}\right)^2 + \sum_{i=1}^{n} \frac{B}{g_i}} = \sum_{i=1}^{n} \frac{c_n - c_i}{g_i} + \sum_{i=1}^{n} \frac{c_n - P}{g_i}. \tag{EC.14}$$

From (EC.12), (EC.13) and (EC.14) it follows that the objective function of problem (P_l^G) simplifies to

$$\sum_{i=l}^{n} \left(\frac{2P - c_i - c_l}{2}\right)\left(\frac{c_i - c_l}{2g_i}\right) + \sum_{i=1}^{n} \frac{(c_n - P)^2}{g_i} - (c_n - P)\sum_{i=l}^{n} \frac{c_i - c_l}{2g_i} \geq 0. \tag{EC.15}$$

Where the inequality follows because the objective function in (EC.15) is quadratic convex in P with a non-positive determinant, for any $0 \le c_1 \le \ldots \le c_n$ and $g_i > 0$ for each $i \in \{1,\ldots,n\}$. Hence, it has non-negative value for any P. Specifically, its determinant is

$$\left(\sum_{i=l}^{n}\frac{c_i-c_l}{g_i}\right)^2+\left(\sum_{i=1}^{n}\frac{1}{g_i}\right)\left(\sum_{i=l}^{n}\frac{(c_i+c_l)(c_i-c_l)}{g_i}\right)-2\left(\sum_{i=l}^{n}\frac{c_i-c_l}{g_i}\right)\left(\sum_{i=1}^{n}\frac{c_n}{g_i}\right)$$
$$\le \left(\sum_{i=l}^{n}\frac{c_i-c_l}{g_i}\right)\left(\sum_{i=l}^{n}\frac{c_i-c_l}{g_i}-\sum_{i=1}^{n}\frac{c_n-c_l}{g_i}\right)\le\left(\sum_{i=l}^{n}\frac{c_i-c_l}{g_i}\right)\left(\sum_{i=l}^{n}\frac{c_i-c_n}{g_i}\right)\le 0.$$

Where the first inequality follows from taking c_n in place of c_i in each term (c_i+c_l) in the second sum, and the second inequality follows from ignoring the terms $i \in \{1,\ldots,l-1\}$ in the last sum.

This completes the proof for the case when constraint (EC.1) is tight. The second case, where constraint (EC.2) is tight instead, is analogous and it is omitted for the sake of brevity. □

EC.3. Proofs of Section 5.2

We formalize the proof sketch of Theorem 3, it follows four steps. First, Proposition EC.8 below allows to write a mathematical program analogous to (P_l^G) in Appendix EC.2, but specialized to the case of Cournot competition with constant marginal costs and linear demand. We denote this problem by (P_l^C), and its optimal objective value by $\mathrm{Z^C}(l,n)$. Moreover, Proposition EC.8 asserts that $\min_{l\in\{1,\ldots,n\}}\mathrm{Z^C}(l,n)$ characterizes a tight bound on the performance of uniform co-payments in maximizing market consumption for this model. Proposition EC.9 asserts that $\mathrm{Z^C}(2,n)$ is at most $\frac{2+\sqrt{2+2/n}}{4}$. Next, Proposition EC.10 asserts that $\mathrm{Z^C}(l,n)$ must be at least $\frac{2+\sqrt{2+2/n}}{4}$, for each $l \in \{1\}\cup\{3,\ldots,n\}$. Hence, we only need to characterize $\mathrm{Z^C}(2,n)$. Finally, the proof of Theorem 3 shows that $\mathrm{Z^C}(2,n) = \frac{2+\sqrt{2+2/n}}{4}$.

Proposition EC.8. *For the special case of $n \ge 2$ Cournot competitors with constant marginal costs $c_i \ge 0$, for each firm $i \in \{1,\ldots,n\}$, facing a linear inverse demand function $P(Q)=a-bQ$, for $a \ge 0$, $b > 0$, and budget $B \ge 0$, let Q^{C} be the market consumption induced by optimal co-payments, and Q^{U} be the market consumption induced by uniform co-payments. Then, the equations in Propositions 2 and 3 can be further specified as follows.*

$$Q^{\mathrm{C}} = \frac{1}{2(m+1)lb}\left((2lm-m+l-1)a-(m+l+1)\sum_{i=1}^{l-1}c_i - lc_l - l\sum_{i=l+1}^{m}c_i\right)$$

$$+\frac{m-l+1}{2(m+1)lb}\left(\left(\left(a+\sum_{i=1}^{l-1}c_i-lc_l\right)^2+\frac{l(m+1)}{m-l+1}\sum_{i=l+1}^{m}(c_i-c_l)^2+\frac{4l(m+1)}{m-l+1}bB\right.\right.$$
$$\left.\left.-\frac{l}{m-l+1}\left(\sum_{i=l+1}^{m}(c_i-c_l)\right)\left(2a+2\sum_{i=1}^{l-1}c_i+\sum_{i=l+1}^{m}c_i-(m+l)c_l\right)\right)^{1/2}. \qquad \text{(EC.16)}$$

$$Q^{\mathrm{U}}=\frac{ua-\sum_{i=1}^{u}c_i+\sqrt{\left(ua-\sum_{i=1}^{u}c_i\right)^2+4u(u+1)bB}}{2(u+1)b}. \qquad \text{(EC.17)}$$

Moreover, there exists an instance of problem (CAP) that minimizes the ratio $Q^{\mathrm{U}}/Q^{\mathrm{C}}$. In particular, any instance of problem (CAP) that minimizes the ratio $Q^{\mathrm{U}}/Q^{\mathrm{C}}$ must be such that $q_n^{\mathrm{U}}=0$. Namely, such that

$$Q^{\mathrm{U}}=\frac{nc_n-\sum_{i=1}^{n}c_i}{b}. \qquad \text{(EC.18)}$$

Additionally, let $l,m,u\in\{1,\dots,n\}$ be indexes as defined in Propositions 2 and 3, respectively. Then, an instance of problem (CAP) in this special case is such that $l\le m=u=n$ if and only if

$$y_l^{\mathrm{C}}\ge 0 \Leftrightarrow bB+\sum_{i=l+1}^{n}\frac{(c_i-c_l)^2}{4}-\sum_{i=l+1}^{n}\frac{c_i-c_l}{4(n+1)}\left(2a+2\sum_{i=1}^{l-1}c_i+\sum_{i=l+1}^{n}c_i-(n+l)c_l\right)\ge 0. \qquad \text{(EC.19)}$$

$$y_l^{\mathrm{C}}\le\frac{c_l-c_{l-1}}{2}\Leftrightarrow\sum_{i=l}^{n}\frac{c_i-c_{l-1}}{4(n+1)}\left(2a+2\sum_{i=1}^{l-1}c_i+\sum_{i=l}^{n}c_i-(n+l)c_{l-1}\right)-\sum_{i=l}^{n}\frac{(c_i-c_{l-1})^2}{4}-bB\ge 0. \qquad \text{(EC.20)}$$

$$q_n^{\mathrm{C}}\ge 0\Leftrightarrow bB+\sum_{i=l+1}^{n}\frac{(c_i-c_l)^2}{4}-\sum_{i=l+1}^{n}\frac{c_i-c_l}{4(n+1)}\left(2a+2\sum_{i=1}^{l-1}c_i+\sum_{i=l+1}^{n}c_i-(n+l)c_l\right)$$
$$+\frac{n-l+1}{4l(n+1)}\left(\left(a+\sum_{i=1}^{l-1}c_i-lc_l\right)^2-\left((n+1)c_n-\sum_{i=1}^{n}c_i-a\right)^2\right)\ge 0. \qquad \text{(EC.21)}$$

$$q_n^{\mathrm{U}}\ge 0\Leftrightarrow bB-\left(nc_n-\sum_{i=1}^{n}c_i\right)\left((n+1)c_n-\sum_{i=1}^{n}c_i-a\right)\ge 0. \qquad \text{(EC.22)}$$

Hence, under the same assumptions of Theorem 3, if $\min_{l\in\{1,\dots,n\}}\mathrm{Z}^{\mathrm{C}}(l,n)=\alpha(n)$ it follows that $\frac{Q^{\mathrm{U}}}{Q^{\mathrm{C}}}\ge\alpha(n)$. Moreover, this bound is tight.

Proof. The proofs of equations (EC.16)-(EC.22) are analogous to the ones in Propositions 2 and EC.2, and are omitted for the sake of brevity.

For any $n \geq 2$ there are finitely many possible combinations of indexes $l, m, u \in \{1, \ldots, n\}$, $l \leq m$. For each combination, $Q^{\mathrm{U}}/Q^{\mathrm{C}}$ has a closed form given by (EC.16) and (EC.17), which is continuous on the parameters of an instance of the problem (CAP) for this special case: the demand parameters a, b, the marginal cost of each firm c_i, and the budget B. Moreover, from Proposition EC.2 it follows that, in order to study the worst-case performance of uniform co-payments, there is no loss of generality in focusing on instances of the problem (CAP) such that the indexes defined in Propositions 2 and 3 satisfy $l \leq u = m = n$. From the discussion above these instances are exactly the ones that satisfy the constraints (EC.19)-(EC.22). Then, from $0 \leq c_1 \leq \ldots \leq c_n \leq a$, $B \geq 0$, $b \geq 0$, together with constraints $(EC.19)$-$(EC.22)$, it follows that the feasible set of problem parameters is compact. Hence, there exists an instance of problem (CAP) that minimizes $Q^{\mathrm{U}}/Q^{\mathrm{C}}$ for each $n \geq 2$.

Consider any instance of problem (CAP) given by $a, b, \mathbf{c}$, and B such that it minimizes the ratio $Q^{\mathrm{U}}/Q^{\mathrm{C}}$. From Lemma EC.2 below it follows that we can assume $c_1 > 0$. Let $(\mathbf{q}^{\mathrm{U}}, Q^{\mathrm{U}})$ be the solution induced by uniform co-payments. Assume for a contradiction that $q_n^{\mathrm{U}} > 0$. Then, we can increase the value of c_n, and reduce the value of c_1, by the same $\epsilon > 0$ sufficiently small, without affecting Q^{U}, while obtaining a strictly larger value for Q^{C}. Specifically, let $(\mathbf{q}^{\mathrm{C}}, Q^{\mathrm{C}})$ be an optimal solution to the original instance of problem (CAP). Consider the modified solution $(\hat{\mathbf{q}}, Q^{\mathrm{C}} + \gamma)$, where $\hat{q}_1 = q_1^{\mathrm{C}} + \delta + \gamma$, $\hat{q}_i = q_i^{\mathrm{C}}$ for each $i \in \{2, \ldots, n-1\}$, and $\hat{q}_n = q_n^{\mathrm{C}} - \delta$, for $\delta > 0$ and $\gamma > 0$ such that $\epsilon = b(\delta + 2\gamma)$, where δ is close enough to $\frac{\epsilon}{b} > 0$, and γ is arbitrarily smaller than δ. This solution is feasible for the modified instance of problem (CAP), and attains an objective value $Q^{\mathrm{C}} + \gamma > Q^{\mathrm{C}}$.

Hence, any instance of problem (CAP) that minimizes $Q^{\mathrm{U}}/Q^{\mathrm{C}}$ must be such that $q_n^{\mathrm{U}} = 0$, or equivalently such that constraint (EC.22) in problem (P_l^C) is an equality. Finally, equation (EC.18) follows from imposing $q_n^{\mathrm{U}} = 0$ in equation (17), for the special case where $g_i = b$ for each $i \in \{1, \ldots, n\}$.

To conclude, note that the equality in the objective function of (P_l^C) follows from equations (EC.16) for Q^{C}, and (EC.18) for Q^{U}, respectively, for the case $u = m = n$, as well as from the definition of $f_l(B,\mathbf{c})$ in (EC.23). Finally, the tightness of the bound is given by the instance of problem (CAP) that attains the minimum at $\min_{l\in\{1,\dots,n\}} \mathrm{Z^C}(l,n)$. $\square$

LEMMA EC.2. *For any $n \geq 2$ Cournot competitors with constant marginal costs $c_i \geq 0$ for each firm $i \in \{1,\dots,n\}$, facing a linear decreasing inverse demand function $P(Q) = a - bQ$, where $a \geq 0$, $b > 0$, and for any budget $B \geq 0$. Consider any instance of the co-payments allocation problem problem (CAP) with $c_1 \geq 0$, and any scaling parameter $\delta \geq 0$, then there exists a modified instance with $c_1 = \delta$ such that the modified instance has the same set of optimal solutions, which attain the same objective value.*

Proof. Consider the modified instance $\hat{a} = (a + \delta - c_1)$, $\hat{\mathbf{c}} = (\mathbf{c} + (\delta - c_1)\mathbf{e})$, where $\mathbf{e}$ is a vector of ones. Any feasible solution in the original instance is feasible in the modified instance, and it attains the same objective value, and viceversa. $\square$

In order to simplify the notation, define the function

$$f_l(a,b,B,\mathbf{c}) \equiv (n-l+1)\left(\left(a+\sum_{i=1}^{l-1} c_i - lc_l\right)^2 + \frac{l(n+1)}{n-l+1}\sum_{i=l+1}^{n}(c_i-c_l)^2 + \frac{4l(n+1)}{n-l+1}bB \right.$$
$$\left. - \frac{l}{n-l+1}\left(\sum_{i=l+1}^{n}(c_i-c_l)\right)\left(2a+2\sum_{i=1}^{l-1}c_i + \sum_{i=l+1}^{n}c_i - (n+l)c_l\right)\right)^{\frac{1}{2}}. \quad \text{(EC.23)}$$

For each number of firms $n \geq 2$, and index $l \in \{1,\dots,n\}$, define the problem (P_l^C) by

$$\mathrm{Z^C}(l,n) \equiv \min_{a,b,B,\mathbf{c}} \frac{Q^{\mathrm{U}}}{Q_l^{\mathrm{C}}} = \frac{2(n+1)l\left(nc_n - \sum_{i=1}^n c_i\right)}{(2ln-n+l-1)a - (n+l+1)\sum_{i=1}^{l-1}c_i - l\sum_{i=l}^{n}c_i + f_l(a,b,B,\mathbf{c})}$$

$$\begin{aligned}
&\text{s.t.} \quad \text{constraints } (EC.19)-(EC.22) \text{ are satisfied} \\
&(\mathrm{P}_l^C) \quad B \geq 0 && \text{(EC.24)}\\
&\qquad a \geq c_n && \text{(EC.25)}\\
&\qquad c_i \geq c_{i-1}, \text{ for each } i \in \{1,\dots,n\} && \text{(EC.26)}\\
&\qquad b \geq 0. && \text{(EC.27)}
\end{aligned}$$

Note that the objective function of problem (P_l^C) is continuous and its feasible set is compact, hence $\mathrm{Z^C}(l,n)$ is well defined for each $n \geq 2$ and $l \in \{1,\dots,n\}$.

Proposition EC.8 suggests a mathematical programming approach to derive a tight bound on the performance of uniform co-payments in this model, denoted by $\alpha(n)$, where n is the number of firms. However, the practicality of this approach hinges on whether we can compute $\min_{l\in\{1,\dots,n\}} \mathrm{Z}^{\mathrm{C}}(l,n)$ for each $n \geq 2$. In order to do this, we will need to compute, or at the very least bound, the value of each $\mathrm{Z}^{\mathrm{C}}(l,n)$ for $l \in \{1,\dots,n\}$. We start by upper bounding $\mathrm{Z}^{\mathrm{C}}(2,n)$ in the next proposition.

PROPOSITION EC.9. *For any $n \geq 2$, $\frac{2+\sqrt{2+2/n}}{4} \geq \mathrm{Z}^{\mathrm{C}}(2,n)$*

Proof. For any $n \geq 2$, and inverse demand parameters $a > 0$, $b > 0$, let $c = \left(\frac{n+\sqrt{\frac{n(n+1)}{2}}}{3n+1}\right) a$, and $B = \left(\frac{(n-1)\sqrt{\frac{n(n+1)}{2}}}{(3n+1)^2}\right)\frac{a^2}{b}$. Finally, let $c_1 = 0$ and $c_i = c$ for each $i \in \{2,\dots,n\}$. It is not hard, but tedious, to show that this instance is feasible for problem (P_2^C), and that it attains an objective value of $\frac{2+\sqrt{2+2/n}}{4}$. We omit the details for the sake brevity. $\square$

We continue bounding the values of $\mathrm{Z}^{\mathrm{C}}(l,n)$ in the next proposition, which relies heavily on the result from Proposition EC.9 above, as well as in several technical results that we state below.

PROPOSITION EC.10. *For any $n \geq 2$, $\mathrm{Z}^{\mathrm{C}}(l,n) \geq \mathrm{Z}^{\mathrm{C}}(2,n)$, for each $l \in \{1,\dots,n\}$.*

Proof. Proposition EC.11 below asserts that any optimal solution to problem (P_1^C) must be such that constraint (EC.19) is tight, or equivalently such that $y_1^{\mathrm{C}} = 0$ (see Proposition EC.8). These instances are also considered when $l = 2$, hence $\mathrm{Z}^{\mathrm{C}}(1,n) \geq \mathrm{Z}^{\mathrm{C}}(2,n)$.

Additionally, for any $n \geq 3$, we must have that $\mathrm{Z}^{\mathrm{C}}(3,n) \geq \frac{2+\sqrt{2+2/n}}{4} \geq \mathrm{Z}^{\mathrm{C}}(2,n)$. Where the first inequality follows from Proposition EC.13 below. The last inequality follows from Proposition EC.9.

Finally, for any $n \geq 4$, and for each $l \in \{4,\dots,n\}$ we have that

$$\mathrm{Z}^{\mathrm{C}}(l,n) \geq \frac{2(nl-n+l-1)}{2nl-n+l-1} \geq \frac{6(n+1)}{7n+3} \geq \frac{2+\sqrt{2+2/n}}{4} \geq \mathrm{Z}^{\mathrm{C}}(2,n). \tag{EC.28}$$

Where the first inequality follows from Proposition EC.12 below, the second inequality follows from the left hand side being increasing in l, and taking $l = 4$. The third inequality holds for any $n \geq 1$, and the last inequality follows from Proposition EC.9. $\square$

We now state and prove several results that where invoked in the proof of Proposition EC.10.

Proposition EC.11. *For any $n \geq 2$, any optimal solution to problem (P_1^C) must be such that $y_1^{\mathrm{C}} = 0$.*

Proof. Note that the problem (P_1^C) can be written as

$$
\begin{aligned}
& \mathrm{Z^C}(1,n) \equiv \\
\min_{a,bB,\mathbf{c}} \quad & \frac{Q^{\mathrm{U}}(\mathbf{c})}{Q_1^{\mathrm{C}}(a,b,B,\mathbf{c})} = \frac{2(n+1)\left(nc_n - \sum_{i=1}^n c_i\right)}{na - \sum_{i=1}^n c_i + f_1(a,b,B,\mathbf{c})} \\
\text{s.t.} \quad & bB \geq \frac{1}{4(n+1)}\left(\sum_{i=2}^n (c_i - c_1)\right)\left(2a + \sum_{i=2}^n c_i - (n+1)c_1\right) - \sum_{i=2}^n \frac{(c_i - c_1)^2}{4} && \text{(EC.29)} \\
(\mathrm{P}_1^C) \quad & bB = \left(nc_n - \sum_{i=1}^n c_i\right)\left((n+1)c_n - \sum_{i=1}^n c_i - a\right) && \text{(EC.30)} \\
& a \geq c_n && \text{(EC.31)} \\
& c_i \geq c_{i-1}, \text{ for each } i \in \{1,\dots,n\} && \text{(EC.32)} \\
& b \geq 0, && \text{(EC.33)}
\end{aligned}
$$

where constraints (EC.29) and (EC.30) correspond to (EC.19) and (EC.22) in the generic problem (P_l^C). From $l=1$ we can drop (EC.20) from (P_l^C), because there is no firm with index $(l-1)$. Moreover, we additionally dropped constraints (EC.21) and (EC.24) from (P_l^C), because (EC.21) is redundant with (EC.29), while $B \geq 0$ is redundant with (EC.30) and (EC.29).

Note that $a=0$ or $b=0$ implies that the objective function of problem (P_1^C) is equal to 1. Hence, we focus on $a>0$, $b>0$. For any fixed $a>0$, $b>0$, let $(B^*, \mathbf{c}^*)$ be an optimal solution to problem (P_1^C). Note that if the k largest variables c_i^* are equal to c_n^*, with $k \in \{1,\dots,n-1\}$, then the objective function is strictly increasing in c_n. It follows that c_n^* must attain its lower bound, otherwise we could strictly improve the objective by decreasing it. We will use this argument iteratively. Specifically, starting from the case $k=1$ it follows that either constraint (EC.29) is tight, or we must have $c_n^* = c_{n-1}^*$. If constraint (EC.29) is tight, we are done. Therefore, assume that $c_n^* = c_{n-1}^*$. In fact, by iterating this argument for each $k \in \{2,\dots,n-2\}$, and conclude that either constraint (EC.29) is tight, or we must have $c_n^* = c_i^*$ for each $i \in \{2,\dots,n\}$. Again, if constraint (EC.29) is tight, we are done. Therefore, assume that $c_n^* = c_i^*$ for each $i \in \{2,\dots,n\}$. Then constraint

(EC.29) simplifies to $c_n^* \geq \dfrac{(3n+1)}{(5n+3)} a > 0$. Finally, considering the case $k = n-1$ it follows that c_n^* must attain its lower bound, hence constraint (EC.29) must be tight. This concludes the proof. □

PROPOSITION EC.12. *For any $n \geq 2$, and for each $l \in \{2, \ldots, n\}$, $\mathrm{Z}^{\mathrm{C}}(l,n) \geq \dfrac{2nl - 2n + 2l - 2}{2nl - n + l - 1}$.*

Proof. The proof structure is the following. We consider a mathematical programming relaxation of problem (P_l^C), denoted by (LBP_l), whose optimal solution provides a lower bound on $\mathrm{Z}^{\mathrm{C}}(l,n)$, for any $n \geq 2$, fixed $a > 0$, $b > 0$, and for each $l \in \{2, \ldots, n\}$. We reformulate this relaxation as a linear program, and we use strong duality to obtain its optimal objective value in closed form, which turns out to be independent of a, b.

We start with an observation that will be useful to construct the relaxation of problem (P_l^C). Note that, if the index l defined in Proposition 2 satisfies $l \geq 2$, then

$$Q^{\mathrm{C}} \leq \frac{2ma - 2\sum_{i=1}^{l-2} c_i - (m-l+3)c_{l-1} - \sum_{i=l}^{m} c_i}{2(m+1)b}. \tag{EC.34}$$

This bound is tight when $y_l^{\mathrm{C}} = \frac{c_l - c_{l-1}}{2}$.

First, note that $a = 0$ or $b = 0$ implies that the objective function of problem (P_l^C) is equal to 1, for any $n \geq 2$ and for each $l \in \{2, \ldots, n\}$. Hence, we focus on the case $a > 0$, $b > 0$. Note that for any $n \geq 2$, and for any fixed $a > 0$, $b > 0$, problem (LBP_l) below is a mathematical programming relaxation of problem (P_l^C).

$$\min_{\mathbf{c}} \; \frac{2(n+1)\left(nc_n - \sum_{i=1}^{n} c_i\right)}{2na - 2\sum_{i=1}^{l-2} c_i - (n-l+3)c_{l-1} - \sum_{i=l}^{n} c_i}$$

$$\text{s.t.} \quad (n+1)c_n - \sum_{i=1}^{n} c_i - a \geq 0. \tag{EC.35}$$

$$(LBP_l) \qquad a \geq c_n \tag{EC.36}$$

$$c_i \geq c_{i-1}, \text{ for each } i \in \{1, \ldots, n\}. \tag{EC.37}$$

Specifically, by substituting the function $Q_l^{\mathrm{C}}(B, \mathbf{c})$ with its upper bound from equation (EC.34), in the objective function of problem (P_l^C), we obtain a mathematical programming relaxation whose objective function does not depend on the budget B. Additionally, we ignore constraints (EC.19), (EC.20) and (EC.21) from problem (P_l^C) altogether. Finally, from $c_n \geq c_i$, for each $i \in \{1,\ldots,n\}$, together with the expression for the budget B in constraint (EC.22) in problem (P_l^C), it follows that constraint (EC.24) from problem (P_l^C) is equivalent to $(n+1)c_n - \sum_{i=1}^n c_i - a \geq 0$. We replace constraint (EC.24) and constraint (EC.22) with this linear inequality. We also drop the variables B and b, as they do not play a role anymore.

The relaxation (LBP_l) is a linear fractional program. Hence, from Charnes and Cooper (1962), it follows that it is equivalent to the following linear program

$$
(LP_l) \qquad \begin{aligned}
\min_{t,\mathbf{x}} \ & nx_n - \sum_{i=1}^n x_i \\
\text{s.}t. \ & 0 \leq x_2 && \text{(EC.38)} \\
& x_i \leq x_{i+1} \text{ for each } i \in \{2,\ldots,n-1\} && \text{(EC.39)} \\
& x_n \leq at && \text{(EC.40)} \\
& (n+1)x_n - \sum_{i=1}^n x_i - at \geq 0. && \text{(EC.41)} \\
& 2nat - 2\sum_{i=1}^{l-2} x_i - (n-l+3)x_{l-1} - \sum_{i=l}^n x_i = 1 && \text{(EC.42)} \\
& t \geq 0. && \text{(EC.43)}
\end{aligned}
$$

Note that $x_i = 0$ for each $i \in \{1,\ldots,l-1\}$, $x_i = \frac{1}{2nl-n+l-1}$ for each $i \in \{l,\ldots,n\}$, $t = \frac{l}{(2nl-n+l-1)a}$ is a primal feasible solution. On the other hand, $\lambda = \frac{(l-1)}{2nl-n+l-1}$, $\gamma = -2n\lambda$, $u_i = (l-i-1)\gamma - (n+l-2i-1)\lambda + l - i - 1$ for each $i \in \{2,\ldots,l-2\}$, $u_{l-1} = 0$, $u_i = -(i+1)\gamma - (n-i)\lambda - i$ for each $i \in \{l,\ldots,n-1\}$, and $u_n = 0$, is a dual feasible solution. Moreover, both solutions attain the same objective value $\frac{(l-1)}{2nl-n+l-1}$. Hence, from strong duality in linear programming, it follows that they are primal and dual optimal, respectively, see for example Bertsimas and Tsitsiklis (1997). Therefore, the associated solution $c_i = 0$ for each $i \in \{1,\ldots,l-1\}$, $c_i = \frac{a}{l}$ for each $i \in \{l,\ldots,n\}$, is optimal for problem (LBP_l), and we conclude $\mathrm{Z}^{\mathrm{C}}(l,n) \geq \frac{2nl-2n+2l-2}{2nl-n+l-1}$ for any $n \geq 2$, and for each $l \in \{2,\ldots,n\}$. $\square$

Proposition EC.13. *For any $n \geq 3$, $\mathrm{Z}^{\mathrm{C}}(3,n) \geq \frac{2+\sqrt{2+2/n}}{4}$.*

Proof. For any $n \geq 3$, note that $a=0$ or $b=0$ implies that both the objective function of problem (P_2^C) and (P_3^C) are equal to 1, and we are done in this case. Therefore, consider any fixed $a>0$, $b>0$, and let $(B^*, \mathbf{c}^*)$ be an optimal solution to problem (P_3^C). Note that if constraint (EC.20) is tight, then it follows that $\mathrm{Z^C}(3,n) \geq \mathrm{Z^C}(2,n)$ and we are done. Similarly, if constraint (EC.19) is tight, then it follows that $\mathrm{Z^C}(3,n) \geq \mathrm{Z^C}(4,n) \geq \mathrm{Z^C}(2,n)$, where the second inequality follows from the case $l=4$ in equation (EC.28), and we are done in this case as well. Hence, assume that constraints (EC.19) and (EC.20) are *loose* for $(B^*, \mathbf{c}^*)$.

Lemma EC.3 below asserts that then $(B^*, \mathbf{c}^*)$ must be such that $c_2^* = c_1^*$, and $c_i^* = c_n^*$, for each $i \in \{3, \ldots, n\}$. Therefore, without loss of generality we focus on solutions with this structure. Moreover, from Lemma EC.2 we assume, without loss of generality, that $c_1^* = 0$. It follows that problem (P_3^C) simplifies to the following one variable optimization problem that we denote by (SMP_3^C),

$$\begin{aligned}
&\min_{c_n} \quad \frac{Q^{\mathrm{U}}(c_n)}{Q_3^{\mathrm{C}}(c_n)} = \frac{12(n+1)c_n}{(5n+2)a - 3(n-2)c_n + \sqrt{(n-2)(3c_n - a)(9(3n+2)c_n - (n-2)a)}} \\
&\text{s.}t. \quad \frac{a}{3} \leq c_n && \text{(EC.44)} \\
&(\mathrm{SMP}_3^C)\ c_n \leq \frac{2(5n+2)a}{9(3n+2)}, && \text{(EC.45)}
\end{aligned}$$

where we have dropped the dependency on the budget B by directly substituting it with the expression from the equality constraint (EC.22). Moreover, constraint (EC.44) is equivalent to constraint (EC.20), and constraint (EC.45) is equivalent to both constraints (EC.19) and (EC.21).

Now we show that for any given $n \geq 3$, and $a>0$, any optimal solution c_n^* to problem (SMP_3^C) must have an objective value of at least $\mathrm{Z^C}(2,n)$. Recall that if at any optimal solution to problem (SMP_3^C) constraints (EC.44) or (EC.45) are tight then we are done. Hence, we focus on values of c_n^* such that $\frac{d\left(Q_3^{\mathrm{U}}(c_n^*)/Q_3^{\mathrm{C}}(c_n^*)\right)}{dc_n} = 0$. After simplifying, this constraint is equivalent to

$$\sqrt{(n-2)(3c_n^* - a)(9(3n+2)c_n^* - (n-2)a)} = 3(n-2)c_n^* - \frac{(n-2)^2}{5n+2}a. \qquad \text{(EC.46)}$$

By substituting expression (EC.46) into the objective function, it follows that any interior stationary point c_n^* must be such that its objective value has the following simplified expression

$\frac{Q_3^{\mathrm{U}}(c_n^*)}{Q_3^{\mathrm{C}}(c_n^*)} = \frac{(5n+2)c_n^*}{2na}$. Furthermore, equation (EC.46) is quadratic in c_n^* and its unique non-negative solution is $c_n^* = \left(\frac{3n+\sqrt{6n(n+1)}}{3(5n+2)}\right) a$. Hence, it follows that any interior stationary point c_n^* attains an objective value of $\frac{3+\sqrt{6+6/n}}{6} \geq \frac{2+\sqrt{2+2/n}}{4} \geq \mathrm{Z}^{\mathrm{C}}(2,n)$, where the first inequality holds for any $n \geq 1$, and the second inequality follows from Proposition EC.9. □

LEMMA EC.3. *For any $n \geq 3$, and fixed $a > 0$, $b > 0$, any optimal solution $(B^*, \boldsymbol{c}^*)$ to problem (P_3^C) for which constraints (EC.19) and (EC.20) are loose must be such that $c_2^* = c_1^*$, and $c_i^* = c_n^*$, for each $i \in \{3, \ldots, n\}$.*

Proof. For any given $n \geq 3$, $a > 0$, $b > 0$, consider any optimal solution $(B^*, \mathbf{c}^*)$ to problem (P_3^C) such that constraints (EC.19) and (EC.20) are loose. Note that, for each index $i \in \{2, \ldots, n-1\}$, it must be the case that either $c_i^* = c_1^*$ or $c_i^* = c_n^*$. Specifically, assume for a contradiction that $i \in \{2, \ldots, n-1\}$ is the largest index such that $c_1^* < c_i^* < c_n^*$. Recall that from Lemma EC.2 it follows that we can assume without loss of generality that $c_1^* = \delta > 0$. Then, note that we can transfer an arbitrarily small $\epsilon > 0$ from c_1^* to c_i^* and strictly improve this solution, while maintaining feasibility for problem (P_3^C), a contradiction. To conclude, note that assuming $l = 3$ implies $c_2^* < c_3^*$, otherwise if $c_2^* = c_3^*$ then firm 2 would get a co-payment whenever firm 3 does, contradicting the definition of the index l in Proposition 2 (see Figure 2). It follows that $(B^*, \mathbf{c}^*)$ must have the structure given in the statement of the lemma. □

We conclude by completing the proof sketch of the main result in Section 5.2. It relies in Proposition EC.14, which is stated and proven below.

Proof of Theorem 3 From Propositions EC.8-EC.10 it follows that it is enough to show that for any $n \geq 2$, $\mathrm{Z}^{\mathrm{C}}(2,n) = \frac{2+\sqrt{2+2/n}}{4}$.

First, note that $a = 0$ or $b = 0$ implies that the objective value of problem (P_2^C) is equal to $1 > \frac{2+\sqrt{2+2/n}}{4} \geq \mathrm{Z}^{\mathrm{C}}(2,n)$, for any $n \geq 2$. Hence, we focus on $a > 0$, $b > 0$. For any $n \geq 2$, and fixed $a > 0$, $b > 0$, Proposition EC.14 below describes a one variable mathematical programming relaxation of problem (P_2^C), denoted by $(\mathrm{RP}_{2,1}^C)$, whose optimal objective value is a lower bound for $\mathrm{Z}^{\mathrm{C}}(2,n)$. We show that the instance given in the proof of Proposition EC.9 is the optimal solution to the relaxation ($\mathrm{RP}_{2,1}^{C}$). Because this instance is in fact feasible for the original problem (P_2^C), it follows that it is optimal for this problem as well.

Proposition EC.14 asserts that the objective function of problem ($\mathrm{RP}_{2,1}^{C}$), $Q_1^{\mathrm{U}}(c_n)/Q_{2,1}^{\mathrm{C}}(c_n)$, is quasiconvex in c_n. We show that the instance from Proposition EC.9 is its unique minimizer. Any interior stationary solution c_n^{C} to problem ($\mathrm{RP}_{2,1}^{C}$) must satisfy $\frac{d\, Q_1^{\mathrm{U}}(c_n^{\mathrm{C}})/Q_{2,1}^{\mathrm{C}}(c_n^{\mathrm{C}})}{dc_n} = 0$. This is equivalent to

$$\sqrt{(n-1)(a-2c_n^{\mathrm{C}})\left((n-1)a-2(5n+3)c_n^{\mathrm{C}}\right)} = \frac{2(n-1)(3n+1)c_n^{\mathrm{C}}-(n-1)^2a}{3n+1}. \tag{EC.47}$$

Equation (EC.47) is quadratic in c_n^{C}, and its unique non-negative solution is $c_n^* = \left(\frac{n+\sqrt{\frac{n(n+1)}{2}}}{3n+1}\right)a \in \left[\frac{a}{2}, a\right]$, where $\left[\frac{a}{2}, a\right]$ is the feasible set of problem ($\mathrm{RP}_{2,1}^{C}$), see Proposition EC.14. Hence, we conclude that this is the unique minimizer of ($\mathrm{RP}_{2,1}^{C}$). Moreover, it is easy to check that this corresponds precisely to the instance given in the proof of Proposition EC.9. From its objective value we conclude $\mathrm{Z}^{\mathrm{C}}(2,n) = \frac{2+\sqrt{2+2/n}}{4}$, completing the proof.

Proposition EC.14. *For any $n \geq 2$, and fixed $a>0$, $b>0$, problem* ($\mathrm{RP}_{2,1}^{C}$) *below is a mathematical programming relaxation of problem* (P_2^C)

$$\min_{c_n} \quad \frac{Q_1^{\mathrm{U}}(c_n)}{Q_{2,1}^{\mathrm{C}}(c_n)} = \frac{4(n+1)c_n}{(3n+1)a-2(n-1)c_n+\sqrt{(n-1)(a-2c_n)\left((n-1)a-2(5n+3)c_n\right)}}$$

$$s.t. \quad \frac{a}{2} \leq c_n \tag{EC.48}$$

$$(\mathrm{RP}_{2,1}^{C}) \quad c_n \leq a. \tag{EC.49}$$

Moreover, its objective value is quasiconvex in c_n.

Proof. First, note that for any $n \geq 2$, and fixed $a>0$, $b>0$, problem (RP_2^C) below is a mathematical programming relaxation of problem (P_2^C).

$$\min_{\mathbf{c}} \quad \frac{Q^{\mathrm{U}}(\mathbf{c})}{Q_2^{\mathrm{C}}(\mathbf{c})}$$

$$\text{s.}t. \quad (n+1)c_n - \sum_{i=1}^{n} c_i - a \geq 0 \tag{EC.50}$$

$$(\mathrm{RP}_2^C) \quad a \geq c_n \tag{EC.51}$$

$$c_i \geq c_{i-1}, \text{ for each } i \in \{1,\dots,n\}. \tag{EC.52}$$

The relaxation is similar to problem (LBP_l) in Proposition EC.12, except that in this case we only relax the feasible set, while we keep the original objective function from problem (P_2^C). Specifically, we ignore constraints (EC.19), (EC.20) and (EC.21) from problem (P_2^C). Additionally, from $c_n \geq c_i$ for each $i \in \{1,\dots,n\}$, together with the expression for the budget B in constraint (EC.22) from problem (P_2^C), it follows that constraint (EC.24) from problem (P_2^C) is equivalent to $(n+1)c_n - \sum_{i=1}^n c_i - a \geq 0$. We replace constraints (EC.24) and (EC.22) from problem (P_2^C) with this linear inequality. Finally, we replace the expression for bB from constraint (EC.22) in the objective function (specifically in $f_2(a,b,B,\mathbf{c})$), making it dependent only in $\mathbf{c}$ for any fixed $a>0$, $b>0$ (see equation (EC.55) below).

Now we show that, without loss of generality, solving problem (RP_2^C) is equivalent to solving one of the following one variable optimization problems, for some index $k \in \{1,\dots,n-1\}$.

$$
\begin{aligned}
& & \min_{c_n} \ & \frac{Q_k^{\mathrm{U}}(c_n)}{Q_{2,k}^{\mathrm{C}}(c_n)} = \frac{4(n+1)kc_n}{(3n+1)a - 2(n-k)c_n + f_{2,k}(c_n)} & \\
(\mathrm{RP}_{2,k}^C) & & \text{s.}t. \ & \frac{a}{k+1} \leq c_n & \text{(EC.53)}\\
& & & c_n \leq a, & \text{(EC.54)}
\end{aligned}
$$

where the function $f_{2,k}(c_n)$ is defined in equation (EC.55) below. Note that for any $n \geq 2$, and fixed $a>0$, $b>0$, any optimal solution $\mathbf{c}^*$ to problem (RP_2^C) must satisfy that there exists an index $k \in \{1,\dots,n-1\}$ such that $c_i^* = c_1^*$, for each $i \in \{1,\dots,k\}$, and $c_i^* = c_n^*$, for each $i \in \{k+1,\dots,n\}$. The proof of this statement is identical to the first part of the proof of Lemma EC.3, and it is omitted for the sake of brevity. It follows that, without loss of generality, we can focus on solutions to problem (RP_2^C) with this special structure, which can be parametrized by the number of firms k with their marginal cost equal to c_1^*. Moreover, from Lemma EC.2, we assume that $c_1^* = 0$. Then, in this case, the function $f_2(a,b,B,\mathbf{c})$ in equation (EC.23) simplifies to

$$f_{2,k}(c_n) \equiv \sqrt{(n-1)\left((n-1)a^2 + 2(k+1)(4nk+n+3k)c_n^2 - 4(2nk+n+k)ac_n\right)}. \qquad \text{(EC.55)}$$

Similarly, Q^{C} and and Q^{U} simplify to $Q_{2,k}^{\mathrm{C}}(c_n) \equiv \frac{(3n+1)a-2(n-k)c_n+f_{2,k}(c_n)}{4(n+1)b}$, and $Q_k^{\mathrm{U}}(c_n) \equiv \dfrac{kc_n}{b}$.

Finally, we show that for any $n \geq 3$, $a > 0$, $b > 0$, and for any index $k \in \{2, \ldots, n-1\}$ (note that we exclude $k = 1$), there is no feasible solution to problem $(\text{RP}^C_{2,k})$ that attains an objective value smaller than $\frac{2+\sqrt{2+2/n}}{4}$. The conclusion then follows from the observation that this lower bound is attained by the candidate instance from Proposition EC.9, which is feasible for problem $(\text{RP}^C_{2,1})$ (i.e. for $k = 1$). Hence, solving the latter must be equivalent to solving the problem (RP^C_2).

Note that the objective function in problem $(\text{RP}^C_{2,k})$ is quasiconvex. Hence, its minimum must be attained either at one of the extremes of the feasible interval $c_n \in \left[\frac{a}{k+1}, a\right]$, or at an interior stationary point. We analyze each one of these cases, and show that none of them attains an objective value smaller than $\frac{2+\sqrt{2+2/n}}{4}$.

(i) If $c_n = \frac{a}{k+1}$, then the objective function of problem $(\text{RP}^C_{2,k})$ evaluates to

$$\frac{4(n+1)k}{(n+1)(3k+1)+\sqrt{(n-1)(n+1)(k-1)(k+1)}} \geq \frac{4k}{3k+1+\sqrt{k^2-1}} \geq \frac{8}{7+\sqrt{3}} \geq \frac{2+\sqrt{2+2/n}}{4}.$$

Where the first inequality follows from the left hand side being decreasing in n, and taking the limit as $n \to \infty$. The second inequality follows from the left hand side increasing in k, for any $k \in \{2, \ldots, n-1\} \geq \sqrt{2}$, $n \geq 3$, and taking $k = 2$. Finally, the last inequality holds for any $n \geq 3$.

(ii) If $c_n = a$, then the objective function of problem $(\text{RP}^C_{2,k})$ evaluates to

$$\frac{4(n+1)k}{n+2k+1+\sqrt{(n-1)\left((8k^2+2k-1)n+6k^2+2k-1\right)}} \geq \frac{4k}{1+\sqrt{8k^2+2k-1}} \geq \frac{4k}{3k+2} \geq \frac{2+\sqrt{2+2/n}}{4}.$$

Where the first inequality follows from the left hand side being decreasing in n, and taking the limit as $n \to \infty$. The second inequality follows from $(3k+1)^2 \geq (8k^2+2k-1)$, for any $k \in \{2, \ldots, n-1\}$, $n \geq 3$. The third inequality follows from the left hand side being increasing in k, and taking $k = 2$; it holds for any $n \geq 1$.

(iii) Any interior stationary solution to problem $(\text{RP}^C_{2,k})$ must satisfy $\frac{d\, Q^{\text{U}}_k(c^*_n)/Q^{\text{C}}_{2,k}(c^*_n)}{dc_n} = 0$. After simplifying, this condition is equivalent to $f_{2,k}(c^*_n) = \frac{2(n-1)(2nk+n+k)c^*_n-(n-1)^2 a}{3n+1}$. Substituting this expression in the objective function of problem $(\text{RP}^C_{2,k})$, it follows that any interior stationary solution must satisfy that the objective function of problem $(\text{RP}^C_{2,k})$ evaluates to

$$\frac{(3n+1)kc^*_n}{2na+n(k-1)c^*_n} \geq \frac{3nk+k}{3nk+n} \geq \frac{6n+2}{7n} \geq \frac{2+\sqrt{2+2/n}}{4}.$$

Where the first inequality follows from the right hand side being increasing in c_n^*, and taking its lower bound $c_n^* = \frac{a}{k+1}$. The second inequality follows from the left hand side being increasing in k, and taking $k = 2$. The last inequality holds for any $n \geq 1$.

This completes the proof $\square$