Working papers series

WP ECON 06.35

A new approach to solve non-regular constrained optimization problems. An application to optimal provision of public inputs

A. Jesús Sánchez (U. Pablo de Olavide)

Diego Martínez (U. Pablo de Olavide y Centro de Estudios Andaluces)

JEL Codes: C6, H21, H3, H41, H43
Keywords: Second Best, numerical analysis, constrained optimization, multi-section, optimal taxation, public input.

Department of Economics

A new approach to solve non-regular constrained optimization problems. An application to optimal provision of public inputs[†]

A. Jesus Sanchez
University Pablo de Olavide

Diego Martinez
Centro de Estudios Andaluces and University Pablo de Olavide

December, 2006 (Preliminary version)

Abstract

This paper describes a new method for solving non-regular constrained optimization problems when standard methodologies do not work properly. Our method (the Rational Iterative Multisection Procedure, RIMP) consists of different stages that can be interpreted as different requirements of precision by obtaining the optimal solution. We have performed an application of RIMP to the case of public inputs provision under two tax settings. We prove that the RIMP and the standard Newton-Raphson (NR) method achieve the same results with regular optimization problems while the RIMP takes advantage over NR when facing non-regular optimization problems.

JEL Classification: C6, H21, H3, H41, H43.
Keywords: numerical analysis, constrained optimization, multisection, optimal taxation, public input.

[†] Parts of this paper have been presented at University Pablo Olavide, the 60th Congress of IIPF, the XII Encuentro de Economía Pública and the 3rd IASC World Conference on Computational Statistics & Data Analysis. We are grateful to G. Fernandez de Cordoba and Esther Barrera for their comments and suggestions. The usual disclaimer applies. A. J. Sanchez the hospitality of the Centro de Estudios Andaluces. Both authors acknowledge the .nancial support by the Spanish Institute for Fiscal Studies and the Spanish Ministry of Science (2003-04028 and 2006-04803). Corresponding author: A.J. Sanchez. University Pablo de Olavide. Ctra. Utrera, km. 1. 41013 Seville. Spain. Phone: 34 954 34 8995. E-mail: ajsanfue@upo.es.

1 Introduction

The use of numerical methods is a standard feature in many areas of scientific research. This is partially due to the need of obtaining unambiguous outcomes in problems in which algebra manipulation does not lead to clear results. In addition, numerically-developed analyses allow to reproduce the characteristics of real situations. In this context, the choice of a determined method for solving optimization problems becomes a crucial decision. In a sense, one could say that there no exists the best method to solve optimization problems. By contrast, different options have to be qualified according to its suitability to the functions to be optimized and the constraints.

This paper introduces a new method for solving non-regular constrained optimization problems which is based on evaluations of the objective function in a multisection of the initial set of possible values, reaching the optimum through an iterative process. We take as a basis the contribution by Casado *et al.* (2000). With the procedure we propose here the level of precision required for the solution becomes a crucial criterion. Starting from an initial set of decision variables, the Rational Iterative Multisection Procedure (hereafter RIMP) selects the compatible values among which the constraints are fulfilled with a determined precision. Subsequent evaluations of the objective function lead to choose the optimal values of decision variables for each level of precision. Moreover, non-optimal solutions (and information to assess how far they are from the local or global optima) can be obtained. In this sense, the path followed by the iterative process towards the global optimum is clearly shown. Although our method achieves the same results at optimum than the Newton-Raphson (hereafter NR) procedure, it provides a wide-ranging set of non-optimal values according to the precision required in order to swell the discussion of results.

In addition, RIMP is unaffected by situations in which NR method does not properly work. For instance, the latter procedure may fail out when the objective function has several (local) solutions or when the starting point is in the neighbourhood of points where the derivative of the objective function is zero. Many of these situations are linked to non-convex problems in economics, such as increasing returns to scale in the production function. Under these circumstances, convergence of NR method towards the global optimum is not guaranteed.

For a better understanding of RIMP, this paper uses the debate around the optimal level of public spending when distortionary taxation is involved. On the basis of the paper by Atkinson and Stern (1974), contributions such as Wilson (1991), Gaube (2000) or Chang (2002) highlight this issue employing in many cases numerical examples (and counterexamples). The underlying idea is that the optimal level of public goods with distorting taxes is below its first-best level.

This debate has not been translated to the case of public inputs. Feehan and Matsumoto (2002), for instance, show the differences between the first and the second-best rules in the provision of public inputs, but nothing is said about the optimal level of public input. By contrast, our paper gives some insights on the levels of public input with distortionary taxation as a first step in this debate. A further discussion on this issue can be found in Martínez and Sanchez (2006), where the distinction between different types of public inputs is focused.

The structure of the paper is as follows. Section 2 explains how the RIMP works with a brief description of the problem to be solved. Section 3 presents an application of RIMP in which a regular and a non-regular constrained optimization problem are solved; a discussion of the results is also included. Finally, section 4 concludes.

2 General description of the methods

In this section we locate the general framework where the problem to be solved can be placed and the two methods used in its resolution as well. Obviously, we focus our attention upon the RIMP given that the NR procedure is a standard well-known method.

2.1 The problem

Let f be the objective function to optimize:

$$\begin{array}{rcc} f : U \times P \subseteq \mathbb{R}^n \times \mathbb{R}^z & \longrightarrow & \mathbb{R} \\ (u,p) & \longrightarrow & f(u,p), \end{array}$$

where f is differentiable[1], U is the set of feasible values for the decision variables (u) (which can be one interval or the union of several), P the set of parameter values fixed throughout all the process (p), n the number of decision variables, and z the number of parameters. Let R be the set of constraints of the problem:

$$\begin{array}{rcc} R : U \times P \subseteq \mathbb{R}^n \times \mathbb{R}^z & \longrightarrow & \mathbb{R}^m \\ (u,p) & \longrightarrow & R(u,p) = (R_1, ..., R_m), \end{array}$$

[1]This regularity is imposed because of NR method. RIMP needs a lower level of regularity as we explain later.

where R is differentiable and m the number of constraints[2].

The problem we are interested in solving is:

$$\begin{cases} \max f(u,p) \\ s.t.: R(u,p) = 0 \\ u \in U,\ p \in P. \end{cases} \tag{1}$$

2.2 Newton-Raphson method

This iterative method has at least two advantages: its high convergence speed and its simple structure. Using the properties of the gradient it is straightforward to achieve the point in which the objective function is maximized. The performance of NR method is simple:

Let $\pounds$ be the function to optimize:

$$\begin{array}{cccc} \pounds : X \subseteq \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ x & \longrightarrow & \pounds(x), \end{array}$$

where $\pounds$ is differentiable. The method is used to solve $\pounds(x) = 0$.[3] Given $x_0 \in X \subseteq \mathbb{R}^n / \exists \nabla \pounds(x_0)^{-1}$, the iterative process has the following steps for each $i > 0$:

1. Evaluate $\nabla \pounds(x_i)$.[4]

2. If $\exists \nabla \pounds(x_i)^{-1}$, we calculate the point to be used in the next iteration: $x_{i+1} = x_i - \pounds(x_i) * \nabla \pounds(x_i)^{-1}$.

3. The stop criterion is defined as follows: given $\epsilon > 0$, if $\|x_{i+1} - x_i\| < \epsilon$, then x_{i+1} is the root of the function; otherwise, the procedure continues until this condition to be satisfied[5].

Further information can be found in the large bibliography existing about this method. However, this widely-used method has some relevant caveats. First, the existence of the solution is guaranteed if the domain of function to be optimized is convex. Second, it is neccessary to have the gradient of function $\pounds$ different to zero; otherwise, the method does not converge. Third, if the objective function has multiple solutions, there exists the risk of jumping

[2] Again, the differentiability asumption is necessary because of NR. This assumption is relaxed later.

[3] The problem (1) has been adapted to this nomenclature using its lagrangian.

[4] The gradient of this function can be obtained analitically or numerically.

[5] There are others possibilities for setting the stopping criterion. For example, ϵ could be defined as $\|\frac{x_{i+1}-x_i}{x_{i+1}}\| < \epsilon$.

from a root near the initial point to other possible solutions, neglecting closer and more accurate solutions. And finally, convergence problems may appear when several local optima are involved in the problem. The procedure we explain next overcomes these caveats.

2.3 Rational Iterative Multisection Procedure (RIMP)

The new method we propose is based on Casado *et al* (2000), and consists of an iterative subdivision of the initial decision variables set. Previously, we have selected the points of the grid that satisfy the constraints with a determined precision in each stage. This process continues until the maximum previously-fixed precision condition is achieved. Whereas Casado *et al* (2000) do not consider different levels of precision (because they use the comparison of values of objective function as stopping criteria), our numerical approach has been adapted to take into account the precision with which the constraints are hold. Hence, the constraints of the problem become specially relevant when RIMP is applied[6].

The term *Rational* is included in the name of the method as a result of two features of the new procedure. First, given that the method will converge better or worse towards the optimal values depending on the initial set of values, it is useful to select the initial range of values with certain rationality. Second, the iterative process followed by RIMP sorts the subsequent initial sets according to the level of precision the problem requires. This reorganization is efficiently made: grouping bordering areas, which allows a more effective search of optimal values. A formal description of RIMP is next.

Definition 1 *Let $f : Y \to Z$ be a function between the partially ordered sets Y and Z. We define f as a monotonic function of degree k if there exists k subsets of Y: $Y_1, ..., Y_k$ ($Y = \bigcup_{i=1}^{k} Y_i, Y_i \cap Y_j = \emptyset$ if $i \neq j$), such that f is monotonic in $Y_i, \forall i = 1, ..., k$.*

Definition 2 *Given the problem (1), $\epsilon > 0$ and the set $W \subseteq U$, let $C(\epsilon, W)$ be the set of compatible values in which the constraints are fullfilled with the precision ϵ, i.e.,*

$$C(\epsilon, W) = \{w \in W \mid \|R(w, p)\|_{\max} < \epsilon\}$$

[6] Consequently, RIMP allows a better treatment of problems where the constraints play an important role, for instance those related to public budgets.

Translating the problem (1) to the new nomenclature, the problem to be solved is:

$$\begin{cases} \max f(c,p) \\ c \in C(\epsilon, U), p \in P \end{cases} \tag{2}$$

The definition of some instrumental but essential parameters is necessary as long as the resolution of the problem consists of using different stages in which a determined precision and a number of subdivisions for each interval is set.

Definition 3 *Let S be the number of stages, then:*

- Precision path*, $E = [E_1, ..., E_S]$, is the vector containing the precision required in the different stages of resolution.*
- Subdivision path*, $D = [D_1, ..., D_S]$, is the vector formed by the number of subdivisions considered for each interval.*

Both variables are interrelated because D_s refers to the number of subdivisions used for achieving the precision E_s in U_s, *i. e.*, the feasible values set for the stage s.

Definition 4 *Let $\tilde{c} \in C(E_S, U_S)$ be the solution to the problem (2), that is, the value which satisfies the condition:*

$$f(\tilde{c}, p) \geqslant f(c, p), \forall c \in C(E_S, U_S) \tag{3}$$

Theorem 5 *Let us consider (1). If R is a monotonic function of degree k, there exists a solution for problem (1) using the Rational Iterative Multisection Procedure.*
Proof. *If R is a monotonic function of degree k, then there exists k intervals $\{I_i\}_{i=1}^{k}$ such that R is monotonic in I_i, $\forall i$. Considering these k subintervals in the first stage, and after adjusting the maximum level of precision to the function R, the solution for problem (1) is achieved.*[7] ■

The implementation of this general procedure has specific characteristics due to computational efficiency reasons which depend on the number of decision variables. The following nomenclature is used to distinguish between them where the confusion may appear; RIMPn will refer to the RIMP method which consider n decision variables. For instance, RIMP2 refers to

[7] A solution could be obtained even if $R > 0$, depending on the desired level of precision. For instance, consider $R \ / \ min(R) = 10^{-5}$.

the method facing a problem with two decision variables. In order to make easier the understanding of the general procedure, the RIMP2 is considered next, that is, $n = 2$ and $c = (c^1, c^2)$.[8] Let us consider I as an arbitrary interval of U_s, *i. e.*, the set of feasible values for the stage s.[9] The resolution for the remaining intervals is analogous to this one. Applying this particular notation to the problem (2) yields the following:

$$\begin{cases} \max f(c, p) \\ c \in C(E_s, I), p \in P \end{cases} \tag{4}$$

With the aim of transforming the continuous problem (4) into a discrete problem, the interval I is subdivided according to the parameter D_s. Hence, we obtain the variables:

$$\begin{aligned} \bar{I}_k &= \max\{c^k | c \in I\} \\ \underline{\mathrm{I}}_k &= \min\{c^k | c \in I\}, \end{aligned}$$

where $k = 1, 2$. Next we consider a particular band witdh for the first decision variable, c^1:

$$H_s^1 = \frac{\bar{I}_1 - \underline{\mathrm{I}}_1}{D_s}.$$

Depending on the problem, it may be useful to set $H_s^2 = H_s^1$ to obtain the same scale in the different decision variables[10]. Finally, the vectors $C^1 = \{c_i^1\}$ and $C^2 = \{c_j^2\}$ are built using the above information:

$$\begin{aligned} c_i^1 =& \underline{\mathrm{I}}_1 + (i-1)H_s^1, i = 1, ..., D_s + 1 \\ c_j^2 =& \underline{\mathrm{I}}_2 + (j-1)H_s^2, j = 1, ..., \frac{\bar{I}_2 - \underline{\mathrm{I}}_2}{H_s^2} + 1 \end{aligned}$$

With these vectors the grid for the interval I_s in this stage is $I_s = C^1 \mathrm{x} C^2$. Using these points all the variables of the problem are evaluated, constraints R included. Thus, for each compatible value of the first decision variable that satisfies the constraints R with a precision E_s, the value of the other which maximizes the objective function f is chosen. In other words, for each c_i^1, the set of *good values* of the other decision variable c_j^2, $G(c_i^1, E_s)$, where the constraints are hold with a precision E_s, is defined. Formally,

[8]RIMP1 is obviously simpler than RIMP2, but it would not show some specific steps we are interested in explain.

[9]$I \subseteq U_s \subseteq \mathbb{R}^2$.

[10]For instance, 10 points for $c^1$and 40 for c^2 when $c^1 \in [0, 1]$ and $c^2 \in [0, 4]$.

$$G(c_i^1, E_s) = \{c_j^2 \in C^2 / \|R(c_i^1, c_j^2)\|_{\max} < E_s\},$$

and grouping the different c_i^1's:

$$G_1(E_s) = \{c_i^1 \in C^1 / G(c_i^1, E_s) \neq \emptyset\}.$$

Using this notation, we find out the solution of problem (4) in the stage s solving the next problem:

$$\begin{array}{c} max(f(c_i^1, G(c_i^1, E_s))) \\ s.a. : c_i^1 \in G_1(E_s) \end{array} \tag{5}$$

This requires to evaluate the objective function at the points satisfying the constraints with the precision required E_s. In addition, this strategy allows to get a ranking of results in the intermediate stages, and shows one of key features of RIMP compared to others numerical methods[11].

Whereas the general procedure has been brieftly described above, several comments are necessary to provide some details on the step from stage s to stage $s+1$. The process must continue searching for values in which the constraints R are hold with the required precision E_{s+1} starting from the discrete set $G_1(E_s)$x$\bigcup_{i=1}^{D_s+1} G(c_i^1, E_s) \subset C(E_s, I_s)$. Thus, for each c_i^1, we form areas around these values in the following way:

- For c_i^1, the RIMP form the interval: $[c^1_{\max\{i-1,1\}}, c^1_{\min\{i+1,D_s+1\}}]$.
- With respect to the second decision variable, the coordinates which belong to $G(c_i^1, E_s)$ are available[12]. Grouping the consecutive numbers obtained, the different areas where RIMP2 will search in the next stage are obtained considering the minimum ($\underline{h}_q^i$) and the maximum ($\bar{h}_q^i$) coordinate of each consecutive subsequence $q = 1, ..., Q_i$, where Q_i is the number of subsequences[13].

Analitically, the process could be summarized as follows: $\forall c_i^1 \in G_1(E_s)$, $\forall q = 1, ..., Q_i$, RIMP2 chooses $[c^2_{\underline{h}_q^i - 1}, c^2_{\bar{h}_q^i + 1}]$, *i. e.*,

$$\bigcup_{c_i^1 \in G_1(E_s)} \bigcup_{q=1}^{Q_i} [c^1_{\max\{i-1,1\}}, c^1_{\min\{i+1,D_s+1\}}] \text{x} [c^2_{\underline{h}_q^i - 1}, c^2_{\bar{h}_q^i + 1}] \subset U_{s+1}$$

[11] The intermediate solutions can be interpreted as solutions of the problem for different levels of precision.

[12] Obviously, this step would not be necessary for RIMP1.

[13] Each consecutive subsequence will form an independent area.

The union of all the intervals created for each interval of U_s will form U_{s+1}, that is,

$$U_{s+1} = \bigcup_{I \in U_s} \bigcup_{c_i^1 \in G_1(E_s)} \bigcup_{q=1}^{Q_i} [c^1_{\max\{i-1,1\}}, c^1_{\min\{i+1,D_s+1\}}] x [c^2_{\underline{\mathrm{h}}^i_q - 1}, c^2_{\overline{h}^i_q + 1}]$$

Finally, the optimal solution will be achieved among the values of this interval by evaluating the objective function. For an intuitive explanation of the transition inter-stages procedure of RIMP, see the *Appendix*.

RIMP has modified our initial problem (1) into a discrete problem. Although RIMP does not invert this process, we can compute how the objective function (and the other variables) is affected when slight changes in the variables involved are presented. The concept of elasticity can be a good illustration of this point. The elasticity of Y with respect to the variable X (e_X^Y) is defined as follows:

$$e_X^Y = (\triangle Y/Y)/(\triangle X/X)$$

Indeed, RIMP can be used to study other issues such as the sensitivity of optimal values to changes in decision variables. Using this concept of elasticity, a comparison of the effects caused by desviations from the optimal values could be carried out. Following the example of elasticity, lower absolute values of e_X^Y imply that the solution achieved is more reliable because there exist less incentives to take different options than the optimal values. RIMP takes advantage here over other methods because this analysis of sensitivity can be done without additional computations.

In the next section, RIMP is used to solve a government problem related to the optimal choice of tax rates and productive public spending. The theoretical model we sketch below is a simple version of that presented in Martínez and Sanchez (2006).

3 An application: the optimal level of public inputs

3.1 The model

We assume an economy of n identical households whose utility function is expressed by $u(x, l)$, where x is a private good used as a numeraire and l the labor supply[14]. Let Y be the total endowment of time such that $h = Y - l$ is the leisure. Output in the economy is produced using labour services and a public input g according to the aggregate production function $F(nl, g)$. This function satisfies the usual assumptions: increasing in its arguments and strictly concave. The type of returns to scale does not matter at the moment, and consequently using the Feehan's (1989) nomenclature, the public input can be treated as firm-augmenting (constant returns to scale in the private factor and the public input combined, creating rents) or as factor-augmenting (constant retuns to the private factor, and therefore scale economies in all inputs). Output can be costlessly used as x or g.

Labour market is perfectly competitive so that the wage rate ω is linked to the marginal productivity of labour:

$$\omega = F_L(nl, g), \tag{6}$$

where firms take g as given. Profits may arise and defined as:

$$\pi = F(nl, g) - nl\omega, \tag{7}$$

which will be completely taxed away by government given their inelastic supply[15].

We distinguish two different tax settings. First, we consider a lump-sum tax T so that the representative household faces the following problem:

$$\begin{aligned} & Max \quad u(x, l) \\ s.t. \quad : \quad & x = \omega l - T, \end{aligned} \tag{8}$$

which yields the labour supply $l(\omega, \omega Y - T)$ and the indirect utility function $V(\omega, \omega Y - T)$. It is to be assumed that $l_\omega \geq 0$.

For later use, we describe some comparative statics of $\omega(g, T, n, Y)$ and $\pi(g, T, n, Y)$[16]:

$$\omega_g = \frac{F_{Lg}}{1 - nF_{LL}l_\omega} > 0 \tag{9}$$

$$\omega_T = \frac{nF_{LL}l_T}{1 - nF_{LL}l_\omega} > 0 \tag{10}$$

[14] The properties of $u(x, l)$ are the standard ones to ensure a well-behaved function: strictly monotone, quasiconcave and twice differentiable.

[15] Pestieau (1976) analyzed how the optimal rule for the provision of public inputs has to be modified when these rents are not taxed away.

[16] Note that variables n and Y are exogenously determined. For the sake of simplicity, we will drop them hereafter as arguments in these functions.

$$\pi_T = -\frac{nlF_{LL}l_T}{1 - nF_{LL}l_\omega} < 0 \quad (11)$$

A second scenario is that using a specific tax on labour τ. Under this tax setting, the consumer's optimization problem could be expressed as:

$$\begin{aligned} & Max \quad u(x,l) \\ s.t. \quad : \quad & x = (\omega - \tau) l \end{aligned} \quad (12)$$

obtaining $l(\omega_N, Y)$ and $V(\omega_N, Y)$, where $\omega_N = \omega - \tau$ is the net wage rate. Again for future reference we derive the following results:

$$\omega_\tau = \frac{-nF_{LL}l_\omega}{1 - nF_{LL}l_\omega} > 0 \quad (13)$$

$$\pi_g = F_g - (nF_{LL}l_\omega + 1) nlF_{Lg} \gtreqless 0 \quad (14)$$

$$\pi_\tau = (1 - \omega_\tau) n^2 lF_{LL}l_\omega < 0 \quad (15)$$

The optimization problem of government in the first-best scenario is as follows:

$$\begin{aligned} & \underset{R}{Max} \quad V(\omega(g), \omega Y - R) \\ & s.t. : g = nR, \end{aligned} \quad (16)$$

where $R = T + \pi(g,T)/n$ is the renevue per person[17].

In the second-best scenario, the optimization problem of government is given by:

$$\begin{aligned} & \underset{R}{Max} \quad V(\omega(g), \omega Y - TEB - R) \\ & s.t. : g = nR, \end{aligned} \quad (17)$$

with $R = \tau l + \pi(g,\tau)/n$ and TEB denoting the total excess burden.

With both tax settings and after some manipulations involving the FOC's of both problems and expressions (9)-(11) and (13)-(15), an important condition for the optimal provision of public inputs is achieved:

$$F_g = 1 \quad (18)$$

[17] It is useful here to consider that rents accrue to consumers before being taxing away by government.

The interpretation is straightforward. The production effects of public input must equal its marginal production cost at optimum. This result is consistent with the extension of the production-efficiency theorem by Diamond and Mirrlees (1971). So far, we know that the provision rules are not affected by the existence of distorting taxes, but anything can be said about the level of public spending and tax rates. Next, we solve numerically the optimization problems (16) and (17) to shed some light on this issue.

3.2 Simulation and results[18]

For the sake of simplicity, we consider a standard Cobb-Douglas utility function widely-used in this literature (Atkinson and Stern, 1974; Wilson, 1991). Specifically,

$$U(x,h) = a\log x + (1-a)\log h \tag{19}$$

where $a \in (0,1)$.

The relevant point in our case comes from the specification of the production function because the different alternatives by defining how the private and public factors enter the production function have notable implications on the debate. The main issue here refers to the return of production function. In particular, whether this function exhibits constant returns to scale in public and private inputs (firm-augmenting public input) or only constant returns to the private factors (factor-augmenting public input) have consequences on the controversy.

First, we study whether the RIMP and the standard NR method achieve the same results when the optimization problem has enough regularity properties (the case of firm-augmenting public inputs). With the aim of study whether the dimension of the problem has any effect on the achieved results, a resolution with the dimensionality reduced is considered[19]. Second, RIMP is used for solving non-regular optimization problems (such a factor-augmenting public inputs) taking advantage over NR.

Initial set of values from which RIMP2 will begin its search is $[0,2]$ for t, and $[100,300]$ for g.[20] Precision requirements we impose for searching

[18] The MATLAB routines used are available at
http://www.upo.es/econ/sanchez_fuentes/docs/research/RIMPv1A.zip

[19] This analysis has been done taking the labour supply as constant under the distorting tax scenario. Then, the dimension of the problem is reduced from two to one.

[20] These intervals are large enough to obtain feasible solutions. Moreover, the largest intervals are defined to include all the initial sets. For further information, see comments on Table 4.

solutions are given by the parameter vector $E = [2, 10^{-1}, 10^{-2}]$.[21] With the aim of achieving this precision, the vector $D = [30, 10, 10]$ will be used to set the number of subdivisions of each interval at the different stages[22]. As RIMP1 is computationally quicker, an additional stage will be considered, *i. e.*, the vectors $D = [30, 10, 10, 10]$ and $E = [1, 10^{-1}, 10^{-2}, 10^{-3}]$ are taken account when g lies within $[100, 300]$.

With respect to the NR method, solutions will be extracted from Martinez and Sanchez (2006), where a standard implementation of this method has been used for solving the same problem.

Firm-augmenting public input

We assume a Cobb-Douglas production function given by $F(nl, g) = (nl)^{\alpha} g^{1-\alpha}$, where $\alpha \in (0, 1)$. This specification creates firm-specific rents. As Pestieau (1976) proved, if these rents are also an argument in the consumer's indirect utility function, the optimal spending condition is not the first-best one; however, recall that our model precisely establishes that all economic rents are taxed away by the government. Indeed, the controversy between the first-best and second-best level of public spending has no sense when the firm-augmenting public input creates rents which are completely taxed by the government. Under this scenario, the analytical solution of our model and its numerical resolution give the intuitive result that the optimal level of productive public spending must be exclusively financed with the economic rents.

In a sense, this situation can be compared to that of Feehan and Batina (2004), in which a (semi)public input is equivalent to a common property resource. A Lindahl pricing system is then the appropriate policy instrument, with a charge on firms (only one in our framework) for their utilization and according to the value of public input's marginal contribution to the firms' profits (Sandmo, 1972). All in all, the complete taxation of rents implies to solve the common problem arising when public input provision is involved. Therefore, the production efficiency condition $F_g = 1$ applies here and, consequently, the numerical solution comes from solving the simultaneous equation

[21]The definition of this vector should take into consideration not only the aim of a good precision per se, but also the issues concerning that there a lot of points near the required precision but not satisfying the restriction. In a sense, a trade-off between precision and number of compatible values appears.

[22]Note that D_1 satisfies the following condition: $H_1^1 = 0.1$ and $H_1^1 = H_1^2$. In the following stages, when $D_s = 10$ is considered, a new decimal at each stage for the decision variables is obtained.

system consisting of the production efficiency condition for g and the government budget constraint, with firms and households solving their respective optimization problems. Particularly, we have considered $a = 0.5$, $\alpha = 0.7$ and $n = 100$ as the set of parameters.

INSERT TABLES 1-3 ABOUT HERE

Several comments can be drawn in viewing Tables 1-3. Table 1 compares the results achieved by using different methods. The coincidence of results is the main conclusion. As the lector can observe, there exists a proper level of consistency between RIMP (both cases, RIMP1 and RIMP2) and NR. In addition, no incentives appear to deviate from the optimal values due to the low values of elasticities obtained.

Table 2 shows the optimal path followed by RIMP1 in the different stages. The optimum achieved for each stage could be interpreted as the best choice according to the required level of precision. The decreasing optimal values obtained for the objective function in the different stages (V_{max}) comes from the existence of a trade-off between the level of precision achieved and the number of compatible values. The more precise results are demanded, the less points satisfy these requirements. This table also shows the advantages of this method over NR. Indeed, RIMP permits to qualify the restrictions have to be faced by governments. For instance, consider a policy-maker forced to provide a level of public spending $g \geq 215$. RIMP gives information (on utility levels, for instance) about the most precise solution under these circumstances.

Table 3 reports, for each interval, its initial point (using subindex *ini*), its final point (using subindex *fin*) and the band width H used in computation for the decision variable. In addition, the compatible values found for this variable (g), the indirect utility (V_{max}), and the government constraint R are reported. If no compatible values are found, the minimum of the government constraint R is showed with the aim of comparing it to the level of precision required for the fulfillment of the constraint. This information can be used relaxing enough the precision requirements to find some compatible values. The intermediate stages of RIMP allow to detect the 'good' areas, $[214.8, 215]$ in the second step, which will be used in the next stages[23]. At the same time, the 'bad' or exhausted areas where the method has finished its searching process are obtained.

INSERT FIGURE 1 ABOUT HERE

[23]Following the formal description of RIMP, the interval $[214.7, 215.1]$ will be considered in order to ensure that potentially close 'good' points are not eliminated.

Figure 1 shows a complementary view of the searching process followed by the RIMP1 and already reported in Tables 2 and 3. Each stage focuses on the search of the 'good areas' detected in the previous one. By contrast, the 'bad' or exhausted areas are not taken into account. Again, the RIMP obtains more precise solutions as the level of precision is increased and simultaneously avoids non-efficient computations.

Factor-augmenting public input

The main difference between the above environment and this of factor-augmenting lies in the assumptions on the returns to scale in the production function. Particularly, we assume again a Cobb-Douglas production function but exhibiting increasing returns in all the inputs (constant returns in labor): $F(nl, g) = nlg^{\beta}$, where $\beta \in (0, 1)$. Under this framework, the debate on the level of public spending in alternative tax settings is reborn. Indeed, the use of lump-sum or distorting taxes are necessary as long as rents are null. The following scenario $a = 0.5$, $\beta = 0.2$ and $n = 100$ is considered as the set of parameters.

Solving the government optimization problem with factor-augmenting public inputs is not as straightforward as before. Indeed, the NR algorithm presents some caveats when non-convex sets of constraints are involved. Note that this is our case because we have increasing returns in the production side of the model. Consequently, there is scope for a method such a RIMP.

INSERT TABLES 4-5 ABOUT HERE

Table 4 reports the results achieved when RIMP2 solves the two different tax settings (first-best *vs.* second-best). This table justifies the use of the term 'Rational' in the name of our method. The very different scale for the tax rate obtained for each tax setting requires a previous knowledge for searching the optimal values. In other words, optimal values only can be found whether the search is done in the proper areas. At the same time, the conclusion from the specific problem solved here is very similar to previous studies: there exists a higher level of provision when the government uses lump-sum taxes. Therefore we are here in line with the mainstream of previous literature in which the level reversal is unusual[24].

Table 5 compares two different implementations of the same problem[25]. It reports the tax rate (t), the level of provision of public input (g), the

[24] For further information, see Martinez and Sanchez (2006).

[25] The dimensionality of the problem has been reduced using that the labour supply is constant when a Cobb-Douglas utility function and distorting taxes are considered.

utility (U) and other relevant variables. The equality of results is the main conclusion and gives more robustness to the achieved results. Low values of elasticities implies that the solution is more reliable because it is an indication that the objective function has enough stability around the optimal values.

4 Concluding remarks

This paper has introduced a new numerical method, the Rational Iterative Multisection Procedure (RIMP), which is able to obtain optimal values of a constrained optimization problem, even when the problem present some non-regular properties. In fact, we have dealt with two optimization problems; a problem with enough regularity properties and a non-convex problem as examples of its application. The method is based on an multisection iterative process of the initial set that evaluates the objetive function, obtaining compatible values of the decision variables under several precision requirements. The more stages are considered, the more precise values are obtained. Moreover, there exists a trade-off between the number of compatible values and the precision requirement imposed.

We have compared this new method to the well-known method of Newton-Raphson, when the problem had enough regularity. One of the main conclusions is the coincidence of the results. In order to carry out this comparison, we have used a simple general equilibrium model with public inputs and taxes on labor. The government has to choose the values of fiscal variables to maximize the per capita utility of representative household. This scenario refers to the case of firm-augmenting public inputs.

Moreover, an optimization problem where non-convex sets of constraints are involved has been considered. This is the case dealing with factor-augmenting public inputs. With the aim of avoiding the problems derived from multiple equilibria and corner solutions, we have used the RIMP, which has relative advantages with respect to the standard NR method under these conditions. Our numerical results are clear: the level of public input in the first-best scenario always exceeds that of the second-best, in line with the mainstream of literature dealing with public inputs. Low values of elasticity of variables with respect to the decision variables of the model show clear indications supporting the idea of stability of solutions achieved.

All in all, RIMP becomes a useful tool for solving constrained optimization problems, in which relaxing constraints is a relevant issue. Under these circumstances, RIMP takes advantage over other methods. An example of this could be problems in which legal or constitutional arrangements imply that the budget constraint has not to be fulfilled *strictu sensu*.

Moreover, other applications of RIMP could be studied in a deeper way. For instance, our procedure may be useful to analyze the sensitivity of calibrated parameters in general equilibrium models.

References

[1] Atkinson, A. B. and Stern, N. H. (1974), Pigou, taxation and public goods, *Review of Economic Studies,* 41, 119-128.

[2] Benninga, S. (1989), *Numerical Techniques in Finance*, MIT Press. London, England.

[3] Casado, L.G. , Csendes, T. and García, I. (2000) A New Multisection Technique in Interval Methods for Global Optimization, *Computing*, 65 (3), 263-269.

[4] Chang, M. (2000), Rules and levels in the provision of public goods: the role of complementarities between public good and taxed commodities, *International Tax and Public Finance,* 7, 83-91.

[5] Diamond, P. A. and Mirrlees, J. A. (1971), Optimal taxation and public production I-II, *American Economic Review*, 61, 8-27 and 261-278.

[6] Feehan, J. P. (1989), Pareto-efficiency with three varieties of public input, *Public Finance/Finances Publiques,* 44, 237-248.

[7] Feehan, J. P. and Matsumoto, M. (2002), Distortionary taxation and optimal public spending on productive activities, *Economic Inquiry*, 40 (1), 60-68.

[8] Feehan, J. P. and Batina. R. G. (2004), Public inputs as common property, 38th Annual Meeting of the Canadian Economics Association, Toronto, Ontario.

[9] Gaube, T. (2000), When do distortionary taxes reduce the optimal supply of public goods?, *Journal of Public Economics* 76, 151-180.

[10] Judd, K.L. (1998), *Numerical Methods in Economics*, MIT Press. Cambridge, MA.

[11] Kolda, T. G., Lewis R. M., and Torczon, V. (2003), Optimization by direct search: new perspectives on some classical and modern methods, *SIAM Review*, 45 (3), 385-482.

[12] Martinez, D. and Sanchez, A.J. (2006), On the optimal level of public inputs, WP ECON 06.34, Universidad Pablo de Olavide, Departamento de Economia.

[13] Neumaier, A. (2004) *Complete Search in Continuous Global Optimization and Constraint Satisfaction*, Acta Numerica 2004 (A. Iserles, ed.), Cambridge University Press.

[14] Pestieau, P. (1976), Public intermediate goods and distortionary taxation, *European Economic Review* 7, 351-357.

[15] Press, W.H., Teukolsky, S.A., Vetterlit W.T. and Flannery, B.P. (1996), *Numerical Recipes in Fortran 90: The Art of Parallel Scientific Computing*, Cambridge University Press, New York.

[16] Sandmo, A. (1972), "Optimality rules for the provision of collective factors of production", *Journal of Public Economics,* 1, 149-157.

[17] Wildasin, D. E. (1986), *Urban Public Finance*, Harwood Academic. Chur, Switzerland.

[18] Wilson, J. D. (1991), "Optimal public good provision with limited lump-sum taxation", *American Economic Review,* 81 (1), 153-166.

A Appendix

An intuitive explanation of the inter-stages procedure of RIMP2, supported by a graphical tool, is presented next. The figure below shows the transition from stage 1 to stage 2. In the first stage, the initial set of two decision variables, $t \times g$, has been discretized as a matrix 4x10.[26] All these points are candidates to be solutions of the optimization problem. Assume the points A, B, C, J and F are the points in which the restriction is fulfilled with the minimum precision E_1. Looking for more precise solutions, the areas $1, 2, 3, 4$ and 5 are built to form the initial set to be considered in the second stage.

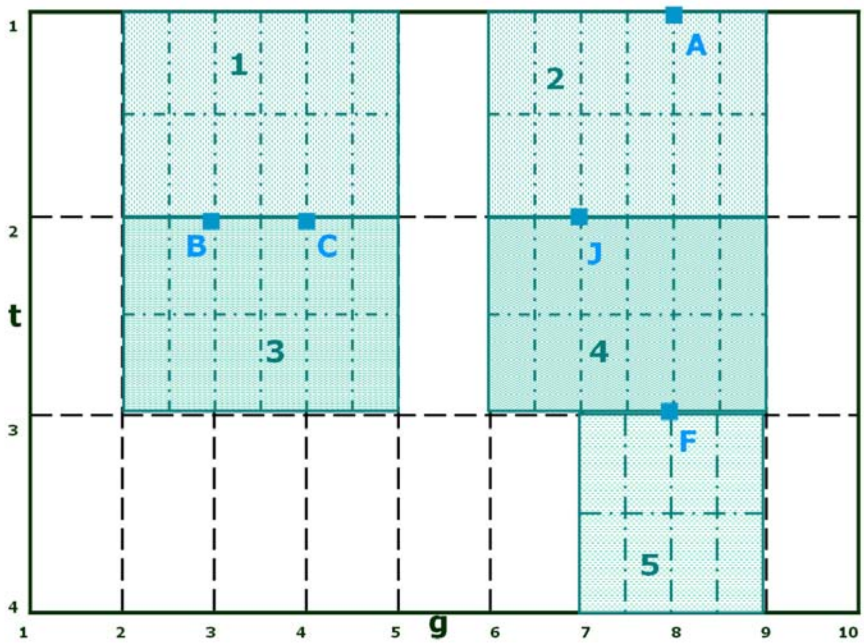

Figure A1: Inter-stages procedure.

The way through which these areas 1-5 are formed can be illustrated taking the coordinate t_2 as a reference. The set of *good values* for the other decision variable g where the constraints are satisfied with a precision E_1 is $G(t_2, E_1) = \{g_3, g_4, g_7\}$. Grouping the consecutive subsequences in g, the minimum and maximum coordinates are, respectively: $\underline{h}_1^2 = 3,\ \bar{h}_1^2 = 4,\ \underline{h}_2^2 = \bar{h}_2^2 = 7$. Hence, the area to be used for the next stage coming from t_2 will be $[t_1, t_3]\text{x}[g_2, g_5] \cup [t_1, t_3]\text{x}[g_6, g_8]$.

However, areas 1-5 are not obtained following *strictu sensu* the theoretical nomenclature explained above or considering directly the area obtained from

[26] For the sake of simplicity, the notation of decision variables is based on the application's nomenclature of section 3.

any of the above points. By contrast, an "efficient" reorganization of the areas relative to each point is done. Therefore, RIMP2 does not duplicate the computation in the areas which belong to more than one point as well as avoids computations in areas where it is unlikely to find good values. Obviously, the total areas are identical in both cases: following the theoretical framework and according to the above efficient reorganization.

Next, the main features of this reorganization are described. We explain them using some particular situations regarding the above figure.

- *The areas where different coordinates are involved should be integrated to optimize the procedure.* For instance, area 2 has been built on the basis of points A and J, which have different coordinates in t. Hence, our method does not duplicate some evaluations corresponding to the common area $[t_1, t_2]$x$[g_7, g_8]$.

- *The good areas found for the interval of coordinates* $[t_{i-1}, t_{i+1}]$ *must be considered in separate intervals.* A good example of this issue could be the situation of the areas 4 and 5. As long as this rule of reorganization would not have taken place, the area $[t_3, t_4]$x$[g_6, g_7]$ would have been included in the second stage and the objective funcion, the constraints and others would have been evaluated in this area, where it is unlikely to find a compatible value in the second stage.

- *The consecutive coordinates in g, which belong to the same coordinate in t, are jointly considered in the definition of the area to be used in the next stage.* An illustration of this situation is given by areas 1 and 3. Points B and C belong to the same subsequence in g, with t_2 as vertical coordinate.

At this point, areas 1-5 are used as the initial set in the second stage and RIMP2 goes on searching for more precise solutions using the new grid of points obtained subdividing these areas according to the parameters D and E.

Figures

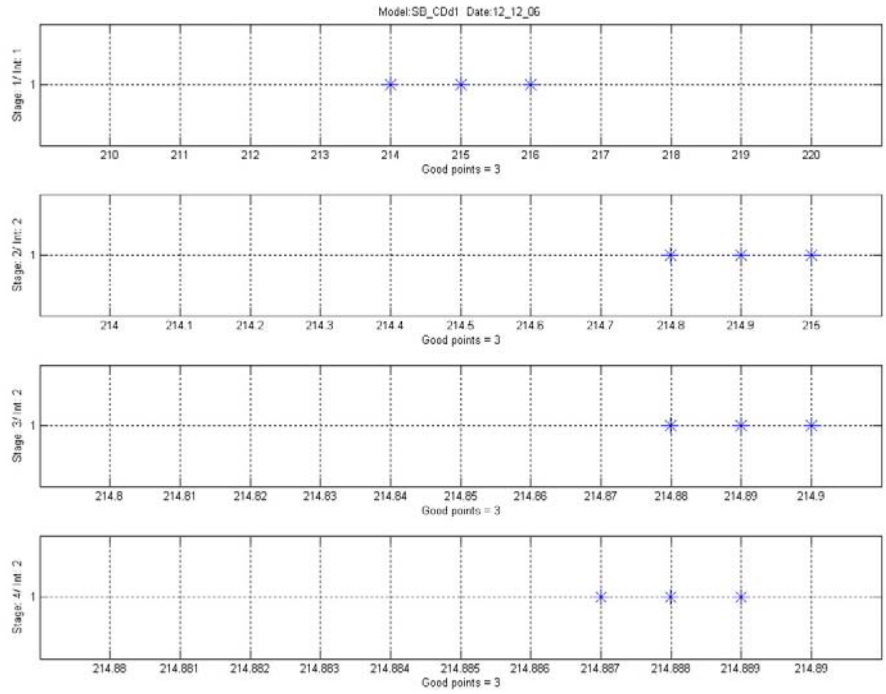

Figure 1: RIMP1 optimal path. Firm-augmenting public input.

Tables

Table 1: Optimal Values. Firm-augmenting public input.

Variable	RIMP1 (D)	e_g^{Var}	RIMP 2 (LS)	e_t^{Var}	e_g^{Var}	NR
Tax Rate (t)	1d-12		1d-12			1d-12
Public Input (g)	214.8890		214.8840			214.8877
Utility (U)	2.0486	0.07322	2.0486	<1E-5	0.07321	2.0486
Labor (l)	12	0	12	<1E-5	0	12
Wages (ω)	0.4178	0.3	0.4178	<1E-5	0.29996	0.4178
Consumption (X)	5.0141	0.3	5.0141	<1E-5	0.29996	5.0140
Total production (F)	716.2930	0.3	716.2888	<1E-5	0.30000	716.2924
Profits (π)	214.8880	0.3	214.8831	<1E-5	0.30010	214.8877
R	8.92E-04		6.00E-04			1.42E-13

R is the precision with which the constraints are satisfied. D distorting. LS = Lump-Sum. $Var = U,\ l, \omega,\ X,\ F,\ \pi$.

Table 2: Optimal path of RIMP1 in different stages. Firm-augmenting public input.

Stage	Interval	g	$V_{\max}$	R	E
1	1	216	2.0493	0.77919	1
2	2	215	2.0487	0.0786	0.1
3	2	214.9	2.0486	0.00859	0.01
4	2	214.89	2.0486	0.00089	0.001

Table 3: Path followed by RIMP1. Firm-augmenting public input.

STAGE = 2 / 4	INTERVAL = 1 / 4	
$g_{ini} = 213$	$g_{fin} = 214$	$H_g = 0.1$
No compatible values found		
Min R = 0.77919 E = 0.1		
STAGE = 2 / 4	INTERVAL = 2 / 4	
$g_{ini} = 214$	$g_{fin} = 215$	$H_g = 0.1$
g	$V_{\max}$	R
214.8	2.04851	0.0614
214.9	2.04858	0.00859
215	2.04865	0.0786
STAGE = 2 / 4	INTERVAL = 3 / 4	
$g_{ini} = 215$	$g_{fin} = 216$	$H_g = 0.1$
g	$V_{\max}$	R
215	2.04865	0.0786
STAGE = 2 / 4	INTERVAL = 4 / 4	
$g_{ini} = 216$	$g_{fin} = 217$	$H_g = 0.1$
No compatible values found		
Min R = 0.77919 E = 0.1		
STAGE = 3 / 4	INTERVAL = 1 / 4	
$g_{ini} = 214.7$	$g_{fin} = 214.8$	$H_g = 0.01$
No compatible values found		
Min R = 0.06140 E = 0.01		
STAGE = 3 / 4	INTERVAL = 2 / 4	
$g_{ini} = 214.8$	$g_{fin} = 214.9$	$H_g = 0.01$
g	$V_{\max}$	R
214.88	2.04857	0.00541
214.89	2.04858	0.00159
214.9	2.04858	0.00859
STAGE = 3 / 4	INTERVAL = 3 / 4	
$g_{ini} = 214.9$	$g_{fin} = 215$	$H_g = 0.01$
g	$V_{\max}$	R
214.9	2.04858	0.00859
STAGE = 3 / 4	INTERVAL = 4 / 4	
$g_{ini} = 215$	$g_{fin} = 215.1$	$H_g = 0.01$
No compatible values found		
Min R = 0.07860 E = 0.01		

Table 4: Optimal path of RIMP2. Factor-augmenting public input.

Distorting						
Stage	Interval	t	g	$V_{\max}$	R	E
1	1	0.11	133.91	2.95353	1.91	2
2	2	0.11	132.1	2.95211	0.1	0.1
3	2	0.11	132.01	2.95204	0.01	0.01
Lump-sum						
Stage	Interval	t	g	$V_{\max}$	R	E
1	1	1.41	142.91	2.95911	0.77919	2
2	15	1.43	143.09	2.95893	0.09568	0.1
3	44	1.43	143.004	2.95886	0.00969	0.01

Table 5: Optimal values with different implementations. Factor-augmenting public input.

Variable	RIMP1	e_g^{Var}	RIMP 2	e_t^{Var}	e_g^{Var}
Tax Rate (t)	0.1100		0.1100		
Public Input (g)	132.0100		132.0100		
Utility (U)	2.9520	0.03534	2.9520	-0.0073	0.03540
Labor (l)	12.0000	<1E-5	12.0000	<1E-5	<1E-5
Wages (ω)	2.6554	0.20000	2.6554	<1E-5	0.20000
Consumption (X)	30.5440	0.20864	30.5442	-0.0432	0.20864
Total production (F)	3186.4200	0.20000	3186.4170	<1E-5	0.20000
Profits (π)	0.0000	-	0.0000	-	-
R	1.00E-02		1.00E-02		

R is the precision with which the constraints are satisfied. $Var = U,\ l, \omega,\ X,\ F,\ \pi$.