GENERALIZED COST-ANALYSIS OF SCREENING PROGRAMS*

Carmen Herrero and Juan D. Moreno-Ternero**

WP-AD 2004-18

Corresponding author: Juan D. Moreno-Ternero. Yale University. Department of Political Science. P.O. Box 208301. New Haven, CT 06520, USA. E-mail: juande@merlin.fae.ua.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

Primera Edición Abril 2004.

Depósito Legal: V-2124-2004

---

* This paper circulated originally under the title "Income-based decision making in the economic evaluation of screening programs". We would like to thank Han Bleichrodt, Mike Drummond, Marc Fleurbaey, Dunia López-Pintado, Andrés Romeu and seminar participants in the Erich-Schneider-Forschungsseminar (Kiel), 2nd Journées D'Economie Publique Louis-André Gérard-Varet (Marseille) and XXVII Simposio de Análisis Económico (Salamanca) for helpful comments and suggestions. Financial support from the *Instituto Valenciano de Investigaciones Económicas* (Ivie) and Fundación BBVA is gratefully acknowledged. Herrero's research was also made possible by project BEC2001-0535 from the Ministerio de Ciencia y Tecnología and by project CTIDIB/2002/314 from the Generalitat Valenciana. Moreno-Ternero's research was also made possible by a postdoctoral fellowship from the Generalitat Valenciana.

** C. Herrero: University of Alicante and Ivie. J.D. Moreno-Ternero: Yale University.

GENERALIZED COST-ANALYSIS OF SCREENING PROGRAMS

Carmen Herrero and Juan D. Moreno-Ternero

ABSTRACT

Nowadays, there is a growing interest about including indirect costs of health care programs in economic evaluations. In this paper, we provide a generalized cost-analysis of screening programs and propose a technique to measure production gains associated with these programs. We apply this technique to show evidence in favor of implementing a newborn screening program to detect congenital hearing impairment.

Keywords: Production gains, screening programs, cost analysis, potential social earnings.

JEL classification numbers: I12, I18.

1 Introduction

Screening has been defined as the systematic application of a test to identify individuals at sufficient risk of a specific disorder to warrant further investigation or direct preventive action, among persons who have not sought medical attention on account of symptoms of that disorder (cf. Jepson et al. 2000). It can be an effective method of reducing morbidity and mortality from disease by detecting it before symptoms occur. On the other hand, it is expensive and can divert resources from other health care programs.

A large medical and epidemiological literature documents controversy over appropriate recommendations about screening for most preventable diseases like prostate cancer, breast cancer or congenital hearing impairment (see Byrne & Thompson (2001), Herrero & Moreno-Ternero (2003), Wu (2003) and the literature cited therein). Among the most important caveats of this literature is the systematic circumvention of the computation of indirect costs associated with screening programs, that might be one of the reasons of the reigning aforementioned controversy. Furthermore, this is in stark contrast with the recent guidelines in *cost-effectiveness analysis* ("*CEA*" hereafter), where a broad agreement seems to be growing to include indirect costs and benefits (e.g., Drummond et al. (1997), Garber (2000), Gold et al. (1996), Olsen & Richardson (1999)).

The aim of this paper is to develop a generalized model that fully describes all sort of costs (direct and indirect ones) associated with screening programs. In particular, we address the computation of costs falling elsewhere in the public sector, the so-called *production costs*. It is well accepted that, in the context of collectively financed health systems, not only costs derived from health care resources consumed but also resources consumed in other sectors should be computed, if society is prepared to consider them in its assessment of health care services. Traditionally, the *human capital method* was used for measuring these costs. Roughly speaking, this method approximates the production costs associated with a disease by the aggregation of the gross earnings an individual would have received if he would have not suffered from the disease. The *friction cost method* (e.g., Brouwer et al. (1997b), Koopmanschap et al. (1995)) is a step beyond in the sense that other realistic situations that are not tackled in the human capital method, like unemployment, are considered to calculate productivity costs. In this paper, we present a method to compute production costs and benefits of screening programs that contemplates a new feature. The feature is that of disentangling the effect of circumstances (aspects that are beyond the control of a person but that affect her pursuit of welfare) and effort (to be understood as those aspects that also influence a person's welfare but that are

not captured by circumstances) in the outcomes of individual earnings. To our viewpoint, the earnings gap due to a disease is well computed only if we compare the earnings of an impaired person with the earnings of a healthy person that has the same remaining circumstances and has expended the same relative degree of effort.

The rest of the paper is organized as follows. In Section 2, we set up the preliminaries. In Section 3, we describe all direct (health service) costs of screening programs. We present in Section 4 our new technique to compute the indirect (non-health service) costs and an elicitation method for it. In Section 5, we apply the model to study the adequacy of implementing a newborn screening program to detect congenital hearing impairment. Finally, we conclude in Section 6.

2 Preliminaries

Let d be a particular disease. Assume that there exist precursors of d such that, if detected before the development of symptoms and treated adequately, might alleviate the consequences of it, even leading to its eradication. Typical examples of d are some congenital impairments (e.g., hearing disorders, phenylketonuria) or some cancers (e.g., prostate or breast cancer). Let $\Gamma = \{1, ..., n\}$ be the corresponding target population of individuals susceptible of suffering the disease d. For instance, if d refers to a congenital disease then Γ typically refers to a cohort of newborns. Individual status with respect to d is either 0 (if the individual is healthy) or 1 (if the individual is impaired).

Every health care policy concerning the fight against d involves two stages: detection and treatment. In the former stage impaired individuals are identified. In the latter one, they are treated and the probability of being cured depends on the early (or late) detection. Early (late) detection occurs when impaired individuals are identified before (after) they develop symptoms of the disease. Usually, early detection is thanks to a screening program, since under conventional management individuals (or their parents, in the case of newborns) consult their general practitioner once they develop symptoms of the disease. Only then, they are referred to a hospital outpatient department for tests and treated if the tests are positive. We assume that the type of management (screening or conventional) does not affect the treatment given. It only affects the probability of a cure.

2.1 Screening

Screening is traditionally defined as testing a population of asymptomatic individuals to identify precursors of a disease. The subjects who test positive are sent on for further evaluation in a subsequent diagnostic test to determine whether they do, in fact, have the disease. Consequently, individuals screened can be partitioned into groups according to whether they do or do not have the disease and whether the screening tests are positive or negative. There are four groups of screened individuals: *true positives*, those whom the screen correctly indicates to have the disease; *false positives*, those who do not have the disease but who have a positive screening test; *false negatives*, those who have the disease but are mistakenly cleared by the screen; and *true negatives*, those who do not have the disease and have a negative screen. As Table 1 shows, the probabilities of an individual being in one of the four groups can be expressed in terms of characteristics of the population (prevalence) and of the detection ability of the screening test (sensitivity and specificity). *Prevalence* is the probability of an individual in the population being impaired, and is measured as $\rho=\frac{TP+FN}{n}$ -the ratio of the number with disease (true positives and false negatives) to the total number of individuals. The *sensitivity* of the screening test (ε) is the ability of the test to identify correctly those who *have* the disease, i.e., the conditional probability that an individual with the disease is positively detected by the test. This is estimated by the ratio of true positives to total impaired individuals ($\frac{TP}{TP+FN}$). The *specificity* of the test (ϕ) is defined as its ability to identify correctly those who *do not have* the disease, i.e., the conditional probability of an individual without the disease being correctly detected as negative in the test. This is measured by the ratio of true negatives to the number of disease-free individuals ($\frac{TN}{FP+TN}$). Using these definitions, the probability of an individual being a true negative is the probability that he does not have the disease $(1-\rho)$ times the probability that the screening correctly indicates that he does not have the disease (ϕ). The probabilities of the individual to be a true positive ($\rho\varepsilon$), a false positive ($(1-\rho)(1-\phi)$) and a false negative ($\rho(1-\varepsilon)$) can be similarly expressed. The advantage of this way of writing the screening probabilities is that it makes easier to assess the implications of variations in the parameters ρ, ε or ϕ separately.

Obviously it would be desirable to have a screening test that is both highly sensitive and highly specific. This is not usually possible, and there is generally a trade-off between the sensitivity and specificity of any given screening test. Screening is often carried out in stages. Normally, a less expensive, less invasive, or less uncomfortable test is carried out first, and those who screen positive on this test are recalled for further testing with

a more expensive or more invasive test, which may have greater sensitivity and specificity. It is hoped that bringing back for further testing those who screen positive will reduce the problem of *false positives.*

Table 1: Impairment status and screening results

Status / Test	Positive	Negative	Total
Positive	TP	FP	$TP+FP$
Negative	FN	TN	$FN+TN$
Total	$TP+FN$	$FP+TN$	$n=TP+FP+FN+TN$

Prevalence: $\rho=\frac{TP+FN}{n}$; Sensitivity: $\varepsilon=\frac{TP}{TP+FN}$; Specificity: $\phi=\frac{TN}{FP+TN}$;

2.2 Conventional management

The alternative to screening is conventional management. Under conventional management individuals who develop symptoms of the disease are referred by their general practitioners to hospital for examination (diagnostic test).

Conventional management therefore defines three groups of individuals; those who have the disease; those who do not but are examined; and those who do not have disease and do not request a test or examination. Following Gravelle et al. (1982), these groups will be referred to as *conventional diseases*, *worried wells* and *unworried wells*, respectively. The probability of an individual being in a particular group defined by conventional management can be expressed in terms of the prevalence ρ and the worried well proportion w (the conditional probability that a well individual develops symptoms is tested and cleared). More precisely, whilst the probability of an individual being a conventional disease is simply the prevalence (ρ), the probability of an individual being a worried well is the probability that he does not have the disease $(1-\rho)$ times the worried well proportion (w). Similarly, the probability of the individual to be a unworried well is $(1-\rho)(1-w)$. We assume that under conventional management all individuals with disease present symptomatically are tested and then treated when the tests are positive.

Table 2: Impairment status and conventional management results

Worried (W) / Status (S)	Positive (W^+)	Negative (W^-)
Positive (S^+)	CD	
Negative (S^-)	WW	UW

3 Health service costs

The effect on national health service costs of screening an individual will be the cost of screening (including the costs of further tests and treatment) less the cost that would have been incurred had he been conventionally managed. The cost of screening an individual will be the sum of the costs associated with each group (true positives, false positives, false positives and true negatives) multiplied by the probability of an individual being in the group. The cost of managing an individual conventionally is the probability weighted sum of the costs of the three groups that define (conventional diseases, worried wells and unworried wells). Thus, in order to measure the effect on health service costs of screening an individual rather than managing him conventionally we estimate the costs associated with the seven groups and their respective probabilities.

Three main types of costs are incurred by the health service: the cost of the screening tests, the cost of the diagnostic tests and the cost of treating the disease. These costs will generally differ for the seven groups defined by the screening and conventional management strategies. Let c_i denote the health care sector (direct) costs associated with each group. Then,

$$c_i = \sigma_i + \delta_i + \tau_i,$$

where σ_i, δ_i and τ_i denote the cost of the screening test, the cost of the diagnostic test and the cost of the treatment in the ith group, respectively. They are described next.

1. **True positives**: Individuals in this group are screened, tested in hospital and treated. Thus, $c_1 = \sigma_1 + \delta_1 + \tau_1$.

2. **False positives**: In this group, they are screened, tested in hospital and cleared, so $c_2 = \sigma_2 + \delta_2$, as $\tau_2 = 0$.

3. **False negatives**: These individuals are screened, mistakenly cleared by the screen, but later during the disease gestation period develop symptoms, are tested in hospital and treated, i.e., $c_3 = \sigma_3 + \delta_3 + \tau_3$. The hospital test and treatment costs might differ from those for true positives because they are incurred later, and so should be discounted. False negatives are also likely to be diagnosed and treated at a later stage of the disease and this might influence treatment costs as well.

4. **True negatives**: They are screened and require no further tests, i.e., $c_4 = \sigma_4$, as $\delta_4 = \tau_4 = 0$.

5. **Conventional diseases**: In this group, individuals are not screened. Once they present symptoms, they are tested in hospital and treated. In other words, $c_5 = \delta_5 + \tau_5$, as $\sigma_5 = 0$. These costs might differ from the hospital test and treatment costs for true positives, as they usually arise later.

6. **Worried wells**: They are not screened but present symptoms through the gestation period and are then tested in hospital and cleared. Then, $c_6 = \delta_6$, as $\sigma_6 = \tau_6 = 0$. The hospital test costs might also differ from those for false positives because of discounting.

7. **Unworried wells**: In this group, there are no health service costs, i.e., $c_7 = 0$.

We assume that the screen test cost is the same for the four groups defined by screening, i.e., $\sigma_i = \sigma$ for all $i = 1, ..., 4$, but permit the hospital and treatment costs to differ across the seven groups. Among them, it is usually assumed, because of discounting and late detection, that

$$\begin{aligned} \delta_1 &= \delta_2 \leq \delta_3 = \delta_5 = \delta_6, \text{ and} \\ \tau_1 &\leq \tau_5 = \tau_3 \end{aligned}$$

Thus, following this notation and the formulae in Section 2, the expected incremental direct costs of introducing a screening, with respect to conventional management are

$$C^D = \rho \cdot (\varepsilon \cdot c_1 + (1 - \varepsilon) \cdot c_3 - c_5) + (1 - \rho) \cdot ((1 - \phi) \cdot c_2 + \phi \cdot c_4 - w \cdot c_6) .$$

4 Non-health service costs

Section 3 concentrates on the effect of screening on national health service costs. We now move to the costs falling elsewhere in the public sector, the so-called production costs. In this section, we present a new method to compute production costs and benefits of screening programs that disentangles the effect of circumstances and effort in the outcomes of individual earnings.

Consider some circumstances for which society should not held responsible individuals in a population. One could think that circumstances are, for instance, the gender, the race, the parental socioeconomic status, the level of formal education attained by parents, and so on. Each individual in the population in question is therefore identified by society with a profile of circumstances. Let $\mathcal{T}$ be the resulting set of types in which the population is partitioned, where a type consists of all individuals who have the same set of

circumstances. Within each type we distinguish those individuals who have the disease d from those who do not. Formally, for all $T \in \mathcal{T}$, $T = T_0 \cup T_1$ where T_0 (T_1) denotes the set of individuals in type T that are healthy (impaired) with respect to the disease d. We denote by ρ_t the prevalence of the disease among individuals sharing the profile of circumstances of type T, i.e., the probability of being in the set T_1.[1]

The population partition $\{T_k : T \in \mathcal{T};\ k \in \{0,1\}\}$ typically changes depending on whether conventional management or a screening program is implemented. Denote by T_k^m (resp. T_k^s) the group of individuals with the profile of circumstances of type T and disease status $k \in \{0,1\}$, after the treatment, when conventional management (resp. a screening program) has been implemented. Then, $\{T_k^m : T \in \mathcal{T};\ k \in \{0,1\}\}$ and $\{T_k^s : T \in \mathcal{T};\ k \in \{0,1\}\}$ are the resulting population partitions after implementing conventional management and a screening program respectively.

Let α_t (resp. α'_t) denote the probability of a cure, for an individual of type T, when the treatment is carried out after an early (resp. late) detection. We may assume, without loss of generality, that $\alpha_t \geq \alpha'_t$ for each $T \in \mathcal{T}$. It is then straightforward to show that the probability of being in the set T_1^m is $\rho_t \cdot (1 - \alpha'_t)$. Similarly, If ρ_t^s denotes the probability of being in T_1^s, and assuming that the sensitivity of the screening program does not depend on the individual circumstances, it is straightforward to show that

$$\rho_t^s = \rho_t \cdot (\varepsilon \cdot (1 - \alpha_t) + (1 - \varepsilon) \cdot (1 - \alpha'_t)), \tag{1}$$

where ε is the sensitivity of the screening program.[2]

4.1 Description of potential social earnings

In order to describe our method, we need to handle earnings distributions. To do so, we follow the so-called *Parade approach* (e.g., Cowell (2000)). This approach can be easily illustrated as follows. Consider two Cartesian axis. Along the horizontal axis, we measure proportions of the population π and earnings e along the vertical axis. The population is arranged in ascending order of earnings and the earnings distribution is given by the pattern composed by all points (π, e_π), where e_π gives the earnings of the person who has exactly a fraction π of the population below her. The usefulness of this approach comes from the fact that statistical concepts can be used and we can compute in an easier way the earnings loss of an impaired individual, that is only attributable to the disease.

[1] Note that $\rho_t = \frac{|T_1|}{n}$, where $|T_1|$ denotes the number of individuals in T_1.

[2] Note that $p_t^m = \frac{|T_1^m|}{n}$ and $p_t^s = \frac{|T_1^s|}{n}$, where $|T_1^m|$ and $|T_1^s|$ denote the number of

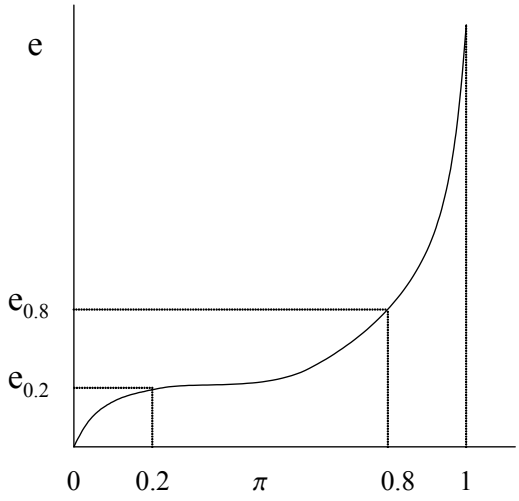

Figure 1: Earnings Distribution

For all $T \in \mathcal{T}$ denote by $m_0^t(\cdot)$ and $m_1^t(\cdot)$ the cumulative distribution functions of the earnings distribution in T_0^m and T_1^m, respectively. Let $x \in T_1^m$ be an impaired individual after conventional management, and e_x denote his earnings. Let π_x be the fraction of the population in T_1^m with lower earnings than him. In other words, x is the π_x^{th} quantile of the earnings distribution in his group.[3] In other words,

$$\pi_x = \int_0^{e_x} dm_1^t.$$

Now, denote by $\overline{e}_x$ the π_x^{th} quantile of the earnings distribution in T_0^m. In other words, $\overline{e}_x$ is the amount of earnings that would leave below in T_0^m the same fraction of the population than e_x in T_1^m, i.e.,

$$\overline{e}_x \text{ is such that } \int_0^{\overline{e}_x} dm_0^t = \pi_x = \int_0^{e_x} dm_1^t.$$

We shall identify the degree of a person's effort with her quantile on the earnings distribution of her group. It is then plausible to think that $\overline{e}_x$ would have been the earnings of individual x in case he would not have suffered the disease, as this only presupposes that the degree of effort expended would be the same in both groups, if the remaining circumstances are identical.

individuals in the set T_1^m and T_1^s respectively.

[3] A quantile is a rank on a distribution normalized to be in the interval $[0, 1]$. It is the infinitesimal version of the discrete analogue of centile.

Thus, the earnings loss of individual x attributable to the disease would be $\overline{e}_x - e_x$. The aggregation of these individual earnings losses is what we call the production loss associated with d, under conventional management. Formally:

Definition 1 *We define the (per capita)* ***production loss*** *associated with d, under conventional management, as*

$$\zeta^m = \frac{1}{n} \cdot \sum_{T \in \mathcal{T}} \sum_{x \in T_1^m} \left(\overline{e}_x - e_x\right).$$

A similar process can be done for the case in which a screening program, rather than conventional management, has been implemented. Formally, for all $T \in \mathcal{T}$ denote by $s_0^t(\cdot)$ and $s_1^t(\cdot)$ the cumulative distribution functions of the earnings distribution in T_0^s and T_1^s. Let $x \in T_1^s$ be an impaired individual after implementing a screening program, and e_x denote his earnings. Let π_x be the fraction of the population in T_1^s with lower earnings than him. In other words, x is the π_x^{th} quantile of the earnings distribution in his group, i.e.,

$$\pi_x = \int_0^{e_x} ds_1^t.$$

Now, denote by $\overline{e}_x$ the π_x^{th} quantile of the earnings distribution in T_0^s. In other words, $\overline{e}_x$ is the amount of earnings that would leave below in T_0^s the same fraction of the population than e_x in T_1^s, i.e.,

$$\overline{e}_x \text{ is such that } \int_0^{\overline{e}_x} ds_0^t = \pi_x = \int_0^{e_x} ds_1^t.$$

Definition 2 *We define the (per capita)* ***production loss*** *associated with d, under a screening program, as*

$$\zeta^s = \frac{1}{n} \cdot \sum_{T \in \mathcal{T}} \sum_{x \in T_1^s} \left(\overline{e}_x - e_x\right).$$

Finally, the *potential social earnings* associated with s are simply given by the difference between both production losses. Formally:

Definition 3 *We define the* ***potential social earnings*** *of a screening program as $\zeta^m - \zeta^s$.*

Thus, the expected incremental indirect costs of introducing a screening, with respect to conventional management, are simply the potential social earnings, with a negative sign, i.e.,

$$C^I = \zeta^s - \zeta^m$$

It is worth noting that the notion of potential social earnings is disentangling the effect of the disease, and other circumstances, with that of effort on individual earnings. In like manner, our proposal is a step beyond the standard human capital method and the friction cost method. It amounts to a more accurate measure of production gains and losses for screening programs.

4.2 A proxy for the elicitation of potential social earnings

The exact computation of potential social earnings would require the follow-up of two comparable populations; one of them under conventional management and the other one under the implementation of a screening program. This might not be feasible in some cases and therefore we ask ourselves which would be the best proxy for that, without the follow-up and when the only data that exist are the actual earnings distributions of the healthy and impaired population, with respect to the disease d. A natural course of action to solve this question is described in the following paragraphs.

We first estimate the current earnings gap that exists between a healthy and an impaired individual, sharing a particular profile of circumstances. The (per capita) production loss associated with d, under conventional management (resp. a screening program) could then be obtained as the weighted aggregation of these gaps, with weights equal to the probability of being in each of the corresponding groups of impaired individuals in the population partition $\{T_1^m : T \in \mathcal{T}\}$ (resp. $\{T_1^s : T \in \mathcal{T}\}$).

Formally, let ξ_t be the estimation of the current existing earnings gap between an impaired individual and a healthy individual, in type T. Then, the estimation of the (per capita) production loss associated with d, under conventional management would be

$$\sum_{T \in \mathcal{T}} \rho_t^m \cdot \xi_t,$$

where, recall that $\rho_t^m = \frac{|T_1^m|}{n}$ denotes the probability of being in the set T_1^m. Similarly, the estimation of the (per capita) production loss associated with d, under a screening program would be

$$\sum_{T \in \mathcal{T}} \rho_t^s \cdot \xi_t,$$

where $\rho_t^s = \frac{|T_1^s|}{n}$ denotes the probability of being in the set T_1^s. Now, making use of (1), it follows that

$$\begin{aligned}\sum_{T\in\mathcal{T}} \rho_t^m \cdot \xi_t - \sum_{T\in\mathcal{T}} \rho_t^s \cdot \xi_t &= \sum_{T\in\mathcal{T}} \xi_t \cdot (\rho_t^m - \rho_t^s) \\ &= \sum_{T\in\mathcal{T}} \xi_t \cdot (\rho_t \cdot (1-\alpha_t') - \rho_t \cdot (\varepsilon \cdot (1-\alpha_t) + (1-\varepsilon)\cdot(1-\alpha_t'))) \\ &= \left(\sum_{T\in\mathcal{T}} \xi_t \cdot \rho_t \cdot (\alpha_t - \alpha_t')\right) \cdot \varepsilon,\end{aligned}$$

where ρ_t is the probability of being in type T, α_t and α_t' are the probabilities of a cure, after early or late detection of the disease, for an individual belonging to type T, and ε is the sensitivity of the screening program.

Thus, we have provided with an outcome measure as a proxy of the potential social earnings that the screening of a particular disease provides. The measure is the following:

$$\left(\sum_{T\in\mathcal{T}} \xi_t \cdot \rho_t \cdot (\alpha_t - \alpha_t')\right) \cdot \varepsilon. \tag{2}$$

It is worth noting that the only screening method-specific information that we need to compute this outcome measure is the sensitivity of the screening program. The outcome measure is indeed a constant factor (that is not screening method-specific) multiplied by the sensitivity of the screening program. Moreover, the constant factor only depends on the estimation of the earnings gap between impaired and healthy individuals (ξ_t), the prevalence within each type (ρ_t) and the probabilities of successful improvement (α_t and α_t'). Without loss of generality, one might assume that $\xi_t \geq 0$. Since we have also assumed that $\alpha_t \geq \alpha_t'$, then the constant factor is non-negative. As a result, and according to this proxy, screening programs for a particular disease could be ranked according to their potential social earnings (indirect costs) just by computing their sensitivity levels. Finally, notice that the value of the proxy depends on the probabilities of improvement, with or without a previous screening: the higher the difference between them, the higher the potential social earnings, as expected.

To conclude with this section, we provide with a concrete expression for ξ_t. A gross one could be the difference between the mean earnings in both subgroups of type T, i.e., $\xi_t = \int_0^\infty e \cdot dt_0(e) - \int_0^\infty e \cdot dt_1(e)$, where $t_0(\cdot)$ and $t_1(\cdot)$ denote the cumulative distribution functions of the earnings distributions in T_0 and T_1. Following the spirit of our model, we propose a more accurate one to reflect the effect of individual effort.

Formally, fix some finite set $\Delta \subset [0,1]$. For each $\pi \in \Delta$, $k \in \{0,1\}$ and $T \in \mathcal{T}$, we denote by $e^{\pi}_{t_k}$ the π^{th} quantile of the earnings distribution in T_k. In other words, $e^{\pi}_{t_k}$ is such that $\pi = \int_0^{e^{\pi}_{t_k}} dt_k$. Then, consider the following estimation of the earnings gap:

$$\xi_t = \frac{1}{|\Delta|} \sum_{\pi \in \Delta} \left(e^{\pi}_{t_0} - e^{\pi}_{t_1} \right), \tag{3}$$

where $|\Delta|$ means the cardinality of Δ.[4]

Finally, it is worth mentioning that the error incurred by proxy (2) in estimating potential social earnings would be zero, under the assumption that the earnings distribution of impaired individuals of each type would be a translation of the earnings distribution of healthy individuals in the same type. This would mean that the existing earnings gap between impaired and healthy individuals, with the remaining circumstances identical, is constant across degrees of relative effort, although it might vary depending on the profile of circumstances. We acknowledge that this assumption is not likely to be fulfilled.

5 Total costs

The guidelines of the US Panel on CEA in health and medicine recommend incorporating indirect costs in the denominator of a *cost-effectiveness ratio* ("*C/E ratio*" hereafter) (cf. Gold et al. (1996)). This recommendation generated a debate between the so-called *"Erasmus group"* and the members of the Panel (cf. Brouwer et al. (1997a), Brouwer et al. (1997b), Weinstein et al. (1997)). The main criticism is that only direct health related effects on quality of life that cannot be meaningfully monetarized should be considered as health effects. Nowadays, it seems to be accepted that production costs and production gains should be included in the numerator of a C/E ratio (cf. Garber (2000), Olsen and Richardson (1999)). Thus, following our model in Sections 3 and 4, the total cost function, or equivalently, the numerator of the C/E ratio, of a screening program could be expressed as follows:

$$C = \rho\left((c_1 - c_3)\varepsilon + (c_3 - c_5)\right) + (1-\rho)\left((c_4 - c_2)\phi + (c_2 - wc_6)\right) + (\zeta^s - \zeta^m).$$

If, moreover, we use proxy (2) to elicit potential social earnings, then we have

$$C = \lambda \cdot \varepsilon + \mu \cdot \phi + \nu, \tag{4}$$

[4] We could also have been more precise by introducing a continuum set of quantiles, $\Delta = [0,1]$. In such a case, $\xi_t = \int_0^1 \left(e^{\pi}_{t_0} - e^{\pi}_{t_1} \right) \cdot d\pi$, for all $T \in \mathcal{T}$.

where

$$\lambda = \rho \cdot (c_1 - c_3) + \sum_{T \in \mathcal{T}} \xi_t \cdot \rho_t \cdot (\alpha_t - \alpha'_t),$$

$$\mu = (1 - \rho) \cdot (c_4 - c_2),$$

and

$$\nu = \rho \cdot (c_3 - c_5) + (1 - \rho) \cdot (c_2 - w \cdot c_6).$$

If (4) is negative, i.e., if screening reduces total costs with respect to conventional management, the work involved in evaluating screening would be much reduced. Although there would be some uncertainty about the precise magnitudes of the reductions in morbidity and mortality arising from earlier detection of a disease, the benefits are almost always certainly positive. Thus if the expression is negative there is no need for a measure of benefit, as screening could be justified solely as a cost reducing innovation.

6 Application: The case of congenital hearing impairment

We conclude by applying our model to the case of congenital hearing impairment. The hearing impairment satisfies all the medical requirements to impose a prevention program, based on a newborn screening protocol. First of all, it is a serious disease, for which a lack of early diagnosis will cause problems in language acquisition. Significant hearing loss interferes with the development of speech perception abilities needed for later language learning. These impairments in communication skills can lead to learning disabilities and ultimately, to limitations in career opportunities. Moreover, it is more frequent than other impairments for which newborn screening programs are in use in developed countries. Finally, there are reliable screening methods, with high levels of sensitivity and specificity, and there is also an effective treatment available (cf. Joint Committee on Infant Hearing (2000)).

In all these arguments that appear in the medical literature, the economic aspect is surprisingly ignored. In this section, we apply our model to show the adequacy of implementing a newborn hearing screening program from an economic perspective. We consider two circumstances: the congenital hearing impairment status itself and the gender. Due to the lack of available data to apply our model directly, we consider the proxy described in Section 4.2 to obtain potential social earnings.

We consider some conservative assumptions in our model that favor conventional management rather than the implementation of a screening program. We shall see that, even in this less favorable framework, there is

enough evidence to support the implementation of a screening program. The assumptions are the following. First, the diagnostic and treatment costs are not increased if they are carried out later. Note that, on average, under conventional management newborns receive diagnostic test and treatment later than under screening management. Second, we assume that the worried wells proportion is zero, i.e., no well individual is tested under conventional management.

Table 3 shows the necessary data for our analysis. In what follows, the reader is referred to Kemper and Downs (2000), Keren et al. (2002) or Kezirian et al. (2001), for additional information about these data. All costs were adjusted to 2001 US dollars. Future costs were discounted at a rate of 3% per year, as recommended by the Panel on Cost-Effectiveness in Health and Medicine (cf. Gold et al. (1996)). The screening cost comprises machine costs, supplies and wages of the specialists. Each machine has a cost which is amortized over 3 years. Supplies include probe tips, probes and machine calibration. We also include here tracking software cost.

Table 3: Data

Parameters	Mean	Range
Prevalence (ρ)	0.0021	$[0.001, 0.006]$
Sensitivity (ε)	0.86	$[0.64, 0.95]$
Specificity (ϕ)	0.97	$[0.93, 1]$
Screening cost (σ)	\$9.52	$[\$7.5, \$12]$
Diagnostic test cost (δ)	\$280	$[\$150, \$540]$
Men's earnings gap (ξ_m)	\$40360	$[\$30010, \$62235]$
Women's earnings gap (ξ_w)	\$99480	$[\$71485, \$111400]$

No satisfactory database including information about earnings distributions within the population of individuals with hearing impairment exists, to the best of our knowledge. For this reason, we used as a proxy the earnings distributions of disabled individuals, which were obtained from the European Community Household Panel (ECHP) relative to Spain. In particular, we obtained from the ECHP the deciles of four earnings distributions: disabled women, disabled men, non-disabled women and non-disabled men. With these data, we applied (3) to obtain an estimation of the earnings gap for both men and women over the whole working period (from 16 to 65 years old). The corresponding values are also shown in the last rows of Table 3.

Finally, we suppose that every impaired infant receives treatment, having a probability α (α') of success if the impairment was early (late) detected. Note that we do not distinguish between these probabilities when they refer to women or men. There is no evidence in the medical literature to reject

this assumption (e.g., Joint Committee on Infant Hearing (2000), Thompson et al. (2001)). Thus, according to (4), the total cost function is given by

$$C = \sigma + (1-\rho)\cdot(1-\phi)\cdot\delta + \rho\cdot(\alpha-\alpha')\cdot\widehat{\xi}\cdot\varepsilon,$$

where $\widehat{\xi} = \frac{1}{2}\cdot(\xi_m + \xi_w)$. Upon replacing the mean values shown in Table 3, it is obtained that if $\alpha - \alpha' \geq 0.142$ then the total cost of implementing a screening program would be negative. That is, if the probability of a cure after early detection is about 15% higher than after late detection, a screening program would save with respect to conventional management.

Univariate sensitivity analyses show that this base-case estimate is quite robust to changes in most of the key parameters and costs. The only two variables to which the model was moderately sensitive were the specificity of the screening program and the prevalence of congenital hearing impairment. More precisely, by varying the specificity within its confidence interval, the threshold for the difference in probabilities would change from 0.075 to 0.23. Similarly, as the prevalence varied within its confidence interval, the threshold changed from 0.05 to 0.298. The ranges of thresholds provided by all the remaining variables are closer to the base-case estimate.

Although there would be some uncertainty about the precise magnitudes of the reductions in morbidity arising from earlier detection of congenital hearing impairment, the benefits are certainly positive and $\alpha - \alpha' \geq 0.142$ is a fairly plausible assumption (e.g., Thompson et al. (2001)). Therefore, the implementation of a newborn hearing screening program could be justified solely as a cost reducing innovation, without requiring the time-consuming task of measuring the benefits associated with these programs. The evidence in favor of a screening program would be even clearer under less conservative assumptions than the ones we considered here. For instance, assuming a non-negative worried wells proportion, or discounting costs of late diagnostic and treatment, the resulting base-case estimate of the threshold for the probabilities would be even lower.

To conclude, it is also worth mentioning that there is a particular feature of this problem that we did not consider. Usually, the treatment of congenital hearing impairment is complemented with a period of special education for impaired infants. It has been argued in the literature (e.g., Grant (2000)) that a screening program provides savings in special education with respect to conventional management. Thus, including this aspect would favor even more the implementation of a screening program.

7 Concluding remarks

In this paper, we have developed a model that fully describes all sort of costs associated with screening programs. In order to compute non-health service costs, we presented a new method that disentangles the effect of effort and the effect of circumstances on individual earnings. We have also applied our model to the particular case of congenital hearing impairment showing that there exists strong support for the implementation of a newborn hearing screening program.

A comment with respect to the computation of non-health service costs (potential social earnings) that we presented is in order. We have considered a utilitarian approach to the aggregation of individual earnings losses. This approach was mainly adopted for simplicity. However, we admit that in so doing, we are implicitly assuming the value judgement that all individual earnings losses are equally valuable. There might be reasons for rejecting this value judgement. For instance, one might argue that rich individuals should be weighted differently than poor individuals in the aggregation of earnings losses. The literature on social choice and welfare provides with several other aggregation methods reflecting some of these value judgements that could be considered. We think that exploring this approach in detail would be beyond the scope of this paper.

To conclude, we acknowledge that our method is inspired by the recent theory of *Equality of Opportunity* developed by Roemer (1998) and it could be interpreted as an application of that theory to a health care context. Roughly speaking, Roemer's theory implies that the effects of the disadvantageous circumstances, beyond the individual's control on his pursuit of welfare, are neutralized, so that the outcome a person eventually achieves is due only to his effort or autonomous choices, where autonomous choice is taken to circumscribe those aspects of a person's behavior which are not determined by circumstance. This implies that individual outcomes may justifiably differ, if they are due only to differential effort or choice, but not if they are due to differential circumstances. In a health care context, equality of opportunity is a conceptualization of equity in the allocation of health care resources. It holds individuals partially responsible for the quality of life-style that they live, in so far as it affects their health, but compensates individuals for the effect on health of circumstances beyond their control. In particular, assume that society implements an equality of opportunity policy à la Roemer and that the objective for which society wants to equalize opportunities is the earning power of individuals. According to this, every impaired individual that has lower earnings just due to a disease for which they should not be held responsible, should be compensated. The aggregation of these compensations

can be seen as a *loss* associated with the disease. This loss could be alleviated by introducing a screening program (instead of conventional management). The magnitude of the reduction in the loss is precisely what we have described as the potential social earnings a screening program offers.

References

[1] Brouwer, W.B.F., Koopmanschap, M.A., & Reuten, F.F.H. (1997a), *Productivity costs measurement through quality of life? A response to the recommendations of the Washington panel*. **Health Economics** 6. 253-259.

[2] Brouwer, W.B.F., Koopmanschap, M.A., & Reuten, F.F.H. (1997b), *Productivity costs in cost-effectiveness analysis: numerator or denominator: a further discussion*. **Health Economics** 6. 511-514.

[3] Byrne, M.M., Thompson, P. (2001), *Screening and preventable illness* **Journal of Health Economics** 20. 1077–1088

[4] Cowell (2000), *Measurement of inequality*, Chapter 2 in **Handbook of income distribution**, A.B. Atkinson and F. Bourguignon (Eds.), Elvesier. New York.

[5] Drummond, M., O'Brien, B.J., Stoddart, G.L. & Torrance, G.W. (1997), **Methods for the economic evaluation of health care programmes.** Second Edition. O.U.P.

[6] Garber, A.M. (2001) *Advances in Cost-effectiveness Analysis of Health Interventions*. Chapter 4 in **Handbook of Health Economics**, A.J. Culyer and J.P. Newhouse (Eds.), North Holland. Amsterdam, The Netherlands.

[7] Grant, R. (2000), *The case to fund universal newborn hearing screening in New York State*. **International Journal of Pediatric Otorhinolaryngology** 54 79–80.

[8] Gravelle H.S.E., P.R. Simpson, J. Chamberlain (1982), *Breast cancer screening and health service costs*, **Journal of Health Economics** 1, 185-207.

[9] Gold M.R., J.E. Siegel, L.B. Russell and M.C. Weinstein (eds.) (1996), **Cost-effectiveness in health and medicine**, Oxford University Press. New York.

[10] Jepson, R., Clegg, A., Forbes, C., Lewis, R., Snowden, A., Kleijnen, J., (2000), *The determinants of screening uptake and interventions for increasing uptake: a systematic review*. **Health Technology Assessment** 4 (14), i–vii, 1–133

[11] Herrero, C., & Moreno-Ternero, J.D. (2002), *Economic evaluation of newborn hearing screening procedures*. **IVIE Working Paper** "A Discusión". WP-AD 2002-06.

[12] Joint Committee on Infant Hearing (2000), *Year 2000 position statement: principles and guidelines for early hearing detection and intervention programs.* **American Journal of Audiology** 9. 9-29.

[13] Kemper, A.R. & Downs, S.M. (2000), *A cost-effectiveness analysis of newborn hearing screening strategies*, **Archives of Pediatrics and Adolescent Medicine** 154 (5).

[14] Kezirian, E.J., White, K.R., Yueh, B. & Sullivan, S. (2001), *Cost and cost-effectiveness of universal screening for hearing loss in newborns*, **Otolaryngol Head Neck Surg.** 124(4). 359-67.

[15] Koopmanschap, M.A., Rutten, F.F.H., van Ineveld, B.M. and van Roijen, L. (1995) *The friction cost method for measuring indirect costs of disease*. **Journal of Health Economics.** 14. 171-189.

[16] Olsen J.A. & Richardson, J. (1999), *Production gains from health care: what should be included in cost-effectiveness analyses?* **Social Science and Medicine.** 49. 17-26

[17] Roemer, J.E. (1998), *Equality of opportunity.* Harvard University Press. Cambridge Ma.

[18] Thompson DC, McPhillips H, Davis RL, Lieu TL, Homer CJ, Helfand M. (2001) *Universal newborn hearing screening: summary of evidence.* **JAMA**. Oct 24-31; 286(16). 2000-10.

[19] Weinstein, M.C., Siegel, J.E., Garber, A.M., Lipscombe, J., Luce, B.R., Manning, W.G., & Torrance, G.W. (1997), *Productivity costs, time costs and health related quality of life: A response to the Erasmus group.* **Health Economics** 6. 505-510.

[20] Wu S. (2003) *Sickness and preventive medical behavior*. **Journal of Health Economics** 22. 675–689.