BAREFOOT AND FOOTLOOSE DOCTORS: OPTIMAL RESOURCE ALLOCATION IN DEVELOPING COUNTRIES WITH MEDICAL MIGRATION

By

John E. Roemer and Pedro Rosa Dias

October 2013
Revised June 2015

COWLES FOUNDATION DISCUSSION PAPER NO. 1922

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

BAREFOOT AND FOOTLOOSE DOCTORS: OPTIMAL RESOURCE ALLOCATION IN DEVELOPING COUNTRIES WITH MEDICAL MIGRATION

JOHN E. ROEMER[1] & PEDRO ROSA DIAS[2]

01.06.15

Abstract

In light of the shortage of healthcare professionals, many developing countries operate a *de facto* two-tiered system of healthcare provision, in which Community Health Workers (CHWs) supplement service provision by fully qualified physicians. CHWs are relatively inexpensive to train but can treat only a limited range of medical conditions. This paper explicitly models a two-tiered structure of healthcare provision and characterizes the optimal allocation of resources between training doctors and CHWs, and implications for population health outcomes. We analyze how medical migration alters resource allocation and population health outcomes, shifting resources towards training CHWs. In the model, migration stimulates health care provision at the lower end of the illness severity spectrum, improving health outcomes for those patients; sufferers of relatively severe medical conditions who can only be treated by doctors are made worse off. It is further shown that donor countries must be reimbursed by *more* than the training cost of emigrating physicians in order to restore *aggregate* population health to the pre-migration level , assuming that there are increasing marginal costs in involved in replacing migrating physicians.

JEL codes: I15, I18 and O15

Keywords: Community health workers, burden of disease, developing countries, migration, optimal resource allocation

Acknowledgements: The authors are grateful to Luigi Siciliani and Alan Winters for their comments and to participants at the *OECD – Universities Joint Conference, 2013* and the conference held in Hermance, Switzerland, in June 2012, organized by the Harvard Program in Ethics and Health and the Brocher Foundation, where an earlier version of this paper was presented. We are grateful as well to two journal referees, who saved us from several errors. Any remaining errors are the authors' responsibility.

---

[1] **Corresponding Author:** Department of Political Science, Yale University, 115 Prospect Street, P.O. Box 208301, New Haven CT 06520-8301; Tel. +12034325249; Email: john.roemer@yale.edu

[2] Department of Economics, University of Sussex, Brighton BN1 9SL, United Kingdom;
Tel. +441273872703; Email: p.rosa-dias@sussex.ac.uk

1. Introduction

The steady increase in life expectancy experienced over the twentieth century is unprecedented in the broad sweep of history. As recently reviewed in Cutler, Deaton and Lleras-Muney (2006), average life expectancy at birth increased by almost 30 years in this period. Nonetheless, a gap of 30 years persists at present between the average life expectancy in rich and poor countries. While this gap is partly attributable to income growth disparities, the bulk of the recent economic literature emphasizes the role of countries' institutional ability and political willingness to adopt and make existing healthcare technologies accessible to populations, as shown in Deaton (2011, 2013). This encompasses both quantity and quality of care provided by health care systems, as shown by recent evidence (for example, Banerjee et al, 2004 and references therein).

This view is broadly consistent with the marked differences in the burden of disease borne by high and low income countries, reflected in the leading causes of mortality shown in Table 1.

Table 1: *Leading causes of death in high and low income countries*

	HIGH INCOME *Fraction total*	**LOW INCOME** *Fraction total*
I. Communicable diseases, maternal and perinatal conditions and nutritional deficiencies	**0.07**	**0.58**
Infectious and parasitic diseases	**0.02**	**0.34**
Tuberculosis	0.00	0.04
HIV/AIDS	0.00	0.08
Diarrhoeal diseases	0.00	0.08
Childhood diseases	0.00	0.01
Malaria	0.00	0.05
Other	0.20	0.07
Respiratory infections	**0.04**	**0.11**
Maternal conditions	**0.00**	**0.02**
Perinatal conditions	**0.00**	**0.09**
Nutritional deficiencies	**0.00**	**0.02**
II. Noncommunicable conditions	**0.87**	**0.33**
Cancer	0.26	0.05
Diabetes mellitus	0.03	0.02
Cardiovascular diseases	0.37	0.16
III. Injuries	**0.06**	**0.09**

Source: WHO, 2008

Infectious and parasitic diseases, such as tuberculosis, diarrheal disease and malaria account for around 35 percent of deaths in low income countries and less than 2 per cent in high income countries. Maternal and perinatal mortality account for more than 10 percent of deaths in low income countries and less than 1 percent of mortality in rich countries. Hence, with the exception of HIV infection, the most prevalent conditions in low income countries are treatable (and some preventable) using existing drugs, treatments and public health interventions, most of them relatively inexpensive and not requiring cutting-edge equipment and infrastructure.

The World Health Report 2006 (WHO, 2006) acknowledges that the provision of relatively simple and inexpensive life-saving interventions, such as antenatal care, immunization and treatment of diarrhea, tuberculosis and malaria, is seriously constrained by a shortage of health workers in the developing world. Table 2 shows the density per 10,000 of population, of fully qualified doctors in Sub-Saharan African countries in 2004.

Table 2: *Density of doctors per 10,000 of population and medical emigration rates in Sub-Saharan Africa*

Country of training	Doctors per 10,000 pop	Doctors emigration (%)
Liberia	0.23	51
Zimbabwe	0.00	45
Ghana	0.90	38
Uganda	0.47	34
South Africa	6.92	34
Malawi	0.11	32
Zambia	0.69	28
Ethiopia	0.29	25
Somalia	0.40	23
Sudan	1.58	19
Tanzania	0.23	15
Rwanda	0.19	14
Nigeria	2.69	13
Togo	0.57	11
Congo, Dem. Rep. of the	0.69	11
Cameroon	0.80	11
Angola	0.80	11
Guinea	0.94	9
Sierra Leone	0.73	9
Congo, Rep. of the	2.51	9
Kenya	1.32	8
Senegal	0.95	6
Mozambique	0.24	6

Source: Data on physician density and emigration used in Bhargava and Docquier (2008, 2012)

With the exception of South Africa, all the countries listed fall considerably short of the minimum requirements set by the WHO (2006) as a pre-requisite for the accomplishment of Millennium Development Goals (MDGs). Although a causality nexus cannot be inferred from the table, it also shows that many of these countries experience high rates of physician emigration, many of them higher than 20% of the total number of fully qualified doctors in the country.

To address this shortage of qualified doctors, many developing countries systematically train and deploy Community Health Workers (CHWs) - a strategy recommended by the WHO (2006). These countries operate a *de facto* two-tiered system of healthcare provision, in which CHWs, recruited from their communities and swiftly trained, supplement service provision by fully qualified physicians. CHW programs have attracted growing attention in the recent economic development literature (see Ashraf, Bandiera and Lee, 2013 and references therein), but have not yet been systematically analyzed in health economics.

We explicitly model a two-tiered structure of healthcare provision, characterize the optimal allocation of resources between training doctors and CHWs and deduce implications for population health outcomes. As shown in Table 2, many of the countries that operate this type of system experience high emigration rates of fully qualified doctors. We analyze how such migration affects resource allocation and population health outcomes, by altering the effective cost of training doctors, thereby shifting resources towards training CHWs. We show that this resource reallocation may benefit patients affected by illnesses treated by CHWs, rendering worse-off sufferers of relatively severe medical conditions, which can only be treated by fully qualified physicians. Finally, we show that donor countries must be reimbursed by more than the training cost of emigrating physicians in order to restore aggregate population health to its pre-migration level. This compensatory payment does not prevent host countries from continuing to benefit from the importation of doctors.

2. Community health workers

The term 'Community Health Worker' is a blanket term used to describe lay members of the community who provide health services, following a short and targeted period of training[3]. China's "barefoot doctors" are arguably the first and most well-known of this type of health worker. Launched in the 1950s, that program aimed at training lay community members to provide primary health care in rural areas, where few qualified doctors wished to settle. The Chinese example spawned a diverse range of healthcare programs throughout the developing world: a (non-exhaustive) list of countries that rely significantly on CHWs for health care provision is given in Table 3 according to a recent WHO report (WHO, 2010).

Table 3: *Alternative designations for CHWs in developing countries*

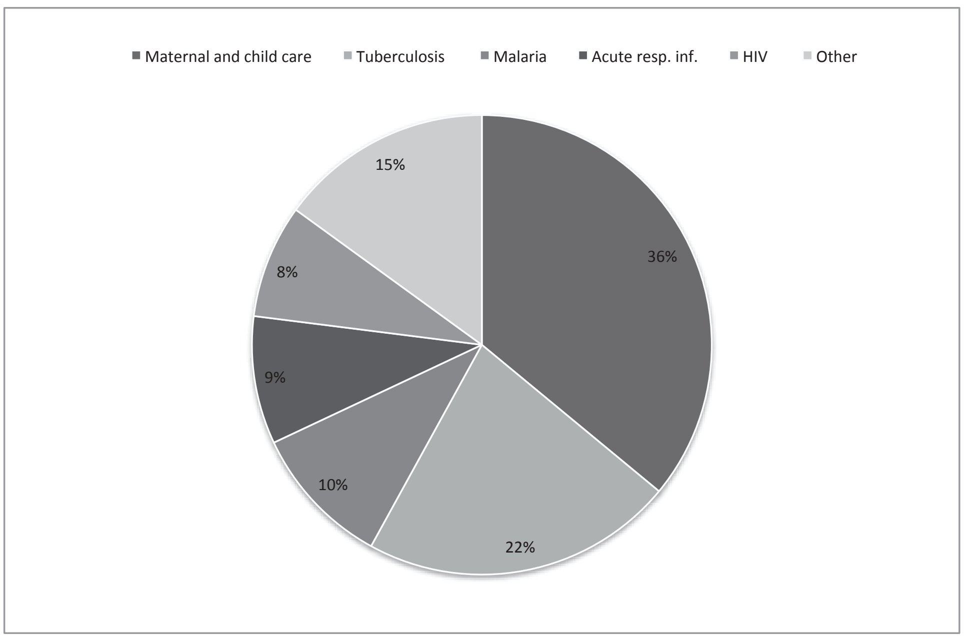
Coutry	CHWs
Bangladesh	Shasthyo Sebika
Peru	Agente Comunitario de Salud
Pakistan	Lady Heath Workers
India	Saksham Sahaya, Maternal & Child Health Promotion Workers amongst others
Brazil	Community Health Agents
Burkina Faso	Women Group Leaders
Burma	Maternal Health Worker
Nepal	Female Community Health Worker
Ethiopia	Village Malaria Worker & Mother Coordinators
Ecuador	Malaria erradication workers
Colombia	Malaria erradication workers
Madagascar	Nutrition workers
Ghana	Nutrition workers
Bolivia	Nutrition workers
Egypt	Raedat
Haiti	Accompagnateurs
Iran	Behvarz
Senegal	Nutrition Worker
Uganda	Community Drug Distributor
Kenya	Village Health Helper
Indonesia	Kader Posyandu
Mali	Village Drug-Kit Manager
South Africa	Lay Health Worker
Uganda	Community Reproductive Health Worker
Guatemala	Village Health Promoters
Nicaragua	Brigadistas

Source: WHO (2010)

[3] According to Lehmann and Sanders (2007), the most widely accepted definition of CHW is the one proposed in WHO (1989): "Community health workers should be members of the community where they work, answerable to the communities for their activities, supported by the health system but not necessarily a part of its organization, and have shorter training than professional workers".

Known by a wide range of country-specific designations, CHWs receive different forms of training and provide different types of care across countries. In most cases, however, their length of training varies from 6 months to two years and is therefore limited to a subset of conditions that a fully qualified doctor can treat. Systematic reviews of healthcare programs based on CHWs, such as WHO (2010) and Lehmann and Sanders (2007), highlight their involvement in outreach activities and curative care in the areas shown in Figure 1. Predictably, these mirror the leading conditions contributing to the burden of disease highlighted in Table 1: a comprehensive range of highly prevalent conditions whose treatment rarely requires a fully qualified medical doctor and complex healthcare technology.

Figure 1. Main activities of CHWs in developing countries



Source: Lehmann and Sanders (2007)

An important characteristic of CHWs is that they are members of the community in which they work. They are thus less likely to attrite and, crucially in our analysis, cannot emigrate and work as health professionals abroad, since their qualifications are not valid outside their country[4].

[4] There is not an international market for CHWs since they are required to have profound knowledge of the local communities, including knowledge of language and culture.

3. Medical migration

Medical migration is frequently identified in the literature as a leading cause of poor health outcomes (Bundred and Levitt, 2000) and short supply of healthcare (for example Ashton et al., 2005) in developing countries. This view is generally shared by health policy officials and officially endorsed by the WHO (World Health Report, 2006 - Chapter 5). It has been rightly argued, as noted in Cutler, Deaton, and Lleras-Muney (2006), that the positive impact of a higher retention of health professionals in donor countries on population health might be hampered by the lack of complementary investments in drugs, equipment and infrastructure that is endemic at present in many developing countries. Nonetheless, a sufficient supply of health professionals remains a fundamental pre-requisite for sizable improvements in health outcomes to be within reach, both in the short and the longer run.

Moreover, the majority of doctors who migrate from developing countries relocate to rich countries. As shown in Table 4, the percentage of doctors trained abroad (most frequently in developing countries) represents roughly one-third of the doctors practicing in the USA, Canada, UK and Australia, and a significant share of other rich countries' medical labor force. Training of medical doctors is expensive even in developing countries: recent estimates in Mills et al (2011) of the cost of fully training a doctor in Sub-Saharan Africa range from $21,000 in Uganda to about $60,000 in South Africa, a cost which, in poor countries, is typically borne by the government. Medical migration therefore implies a partial loss of human capital investment, which is transferred free of charge to the host country, a situation often deemed unfair, as argued in the World Health Report, 2006 – p.101: *(…) when large numbers of doctors and nurses leave, the countries that financed their education lose a return on their investment and end up unwillingly providing the wealthy countries to which their health personnel have migrated with a kind of "perverse subsidy"*

Table 4: *Doctors trained abroad as percentage of practicing doctors, in select OECD countries*

Australia	21
Canada	23
Finland	9
France	6
Germany	6
New Zealand	34
United Kingdom	33
United States	27

Source: WHO (2006)

3.1 Theoretical models

The idea that the attraction of the scarce skilled labor force of developing countries by rich countries is fundamentally unfair has been the focus of the theoretical literature on the *brain drain* since the late 1960s. As reviewed in Commander, Kangasniemi and Winters (2002) and, more recently, in Docquier and Rapoport (2012), the 1970s theoretical literature examines the welfare implications of this brain drain and emphasizes its detrimental effect on developing countries. Stylized models of labor market integration developed in Bhagwati and Hamada (1974, 1975), Rodriguez (1975) and McCulloch and Yellen (1977) indicate that, in the presence of labor market rigidities, imperfect information, externalities and subsidized education in developing countries, emigration of skilled workers affects developing countries negatively. It hinders human capital formation, imposes on them important fiscal costs associated with public provision of education, and, under specific circumstances, may further lead to an increase in unemployment[5]. In order to compensate developing countries for these negative effects, Bhagwati and Hamada (1974) proposed an income tax paid by skilled emigrants – the much discussed *Bhagwati tax*. This would be paid over and above their income tax in the host country and the corresponding tax revenue transferred to the donor country. The debate on the consequences of this proposal, as well as the relative merits of different variations on the *Bhagwati tax* is ongoing[6].

A more recent wave of theoretical models pioneered in the late 1990's re-examines the issue of brain drain in the context of dynamic models and proposes that migration may provide

5 Bhagwati and Hamada (1974) show that, under certain conditions, the possibility of emigration lead the skilled workers of developing countries to bargain for higher wages, leading to an increase of unemployment.

6 See, for example, McCulloch and Yellen (1975, 1977) and, more recently, McHale (2009).

significant positive incentives for skill formation, which might, in net terms, mitigate or even outweigh the loss of human capital that occurs through emigration. Amongst the seminal contributions to this line of research are Mountford (1997), Vidal (1998) and Beine, Docquier and Rapoport (2001). Further theoretical contributions have emphasized additional possible benefits, neglected by the earlier literature, such as migrants' remittances, which are a source of development finance in developed countries. This literature is thoroughly reviewed in Docquier and Rapoport (2012).

The theoretical literature to date has thus largely focused on characterizing the implications of migration of highly qualified individuals on those left behind and on human capital formation dynamics in the donor country, which in turn can impact on its economic growth. In this paper, we abstract from the question of how medical migration impacts on wages of remaining medical personnel in the donor country, and focus instead on its impact on health outcomes, both directly through the loss of qualified medical personnel, and indirectly, through the endogenous shift in the allocation of public spending for the training of medical personnel. To address this issue we explicitly model how both the level of expenditure on health professional training and its allocation across CHWs and physicians impacts on health outcomes, both in the aggregate and along the health status distribution. For parsimony and in line with the first wave of literature on the brain drain, the proportion of emigrating physicians is treated as exogenous.

3.2 Evidence

There is an extensive empirical literature on the overall effect of medical migration on population health outcomes in donor countries. As data are relatively scarce, causal effects are hard to establish and evidence is mixed. Despite these limitations, some associations are well established. Chauvet, Gubert and Mesplé-Somps (2008) show that medical migration is associated with a worsening of child health outcomes in a panel of 98 host countries; interestingly, their results further suggest that medical brain drain reduces the effectiveness of foreign health aid to these countries. Bhargava and Docquier (2008) corroborate the existence of a negative association between the migration of doctors and key population outcomes: doubling the rate of expatriation of fully qualified doctors is associated with a 20 percent increase in adult deaths from AIDS.

While we focus on a distinct new channel through which medical migration can have a detrimental effect on aggregate population health, the economic literature has also emphasized channels for potential gain from migration, as mentioned above. First, medical migration generates remittances: Kangasniemi, Winters and Commander (2007) examine a survey of overseas doctors practicing in the UK in 2002[7]; on average 45% of these doctors sent remittances, on the order of 16% of their earnings, to their families in the donor country. Although remittances may represent a significant source of income for some families in donor countries, they do not directly improve health care quality, availability and population health. Hence, we abstract from these in our model.

Second, doctors who emigrate may return with potentially valuable skills. Although international data are insufficient for a rigorous assessment, Kangasniemi, Winters and Commander (2007) provide useful evidence. Of the migrant doctors who reported to have the intention to return to their donor countries, roughly 65% intended to work in the private sector and almost 90% in urban areas. Returnees are thus unlikely to populate the most impoverished areas in need of care. Moreover, given that 70 per cent of the burden of disease in low-income countries is amenable to simple interventions, the relevance of newly acquired skills in rich countries has been called into question. Overall, the evidence on the hypothesized benefits of a return of doctors who emigrate is, at best, weak. Thus, for simplicity, we abstract from this possibility in our model.

Finally, the recent empirical literature focuses on the plausibility theoretical possibility of 'brain gain', in the sense that the prospect of migration may increase incentives to obtain education, thereby improving, rather than depleting, the stock of human capital. Our model analyses resource allocation in the context of a fairly limited health budget, where the state selects how many doctors to train from a sufficiently large, homogeneous pool of potential candidates. We abstract from the possibility of a shortage of possible individuals to train, assuming health-budget constraints bind. This assumption is plausible given the financial

---

[7] The main donor countries represented in the sample were India (around 42%), Nigeria (8%) and South Africa (roughly 7%). Other Sub-Saharan Africa countries were also represented.

constraints in developing countries, and because the empirical evidence suggests brain gain effects are too small to affect the national stock of doctors[8].

4. The model

This section presents a model of disease and optimal resource allocation to treat it using two types of medical personnel, with and without medical migration. It aims at analyzing the effect of medical migration on health outcomes, both directly through the loss of qualified medical personnel, and indirectly, through the endogenous shift in the allocation of public spending for the training of medics. This sheds light on the relative strength of the mechanisms at play in different scenarios and the quantification of possible compensatory measures. It can also be a basis for empirical work, once more detailed data on CHWs is compiled and made available[9].

We explicitly adopt a social planner's perspective as we consider it to be particularly relevant in our context. In most developing countries, the overwhelming majority of medical training is funded and most often provided by the state: according to a large recent survey funded by the Gates foundation - Mullan et. al. (2010) – in Sub-Saharan Africa over 80 percent of medical school of all kinds are public and their curricula decided in light of local health issues and priorities. Moreover, attendance at the minority of private schools operating in the region is often heavily subsidized by the state. Mullan et al. (2010) also suggests that, in most countries, budget constraints are clearly binding and constitute a severe obstacle to scaling-up health professionals' education[10]. Finally, the adoption a social planner's perspective is

---

[8] Mountford (1997) shows that brain gain hinges on two crucial premises: that migration prospects determine decisions to enroll into medical school and that migrants are not strongly screened by the host country. Kangasniemi, Winters and Commander (2007) find that, for medical migration towards the UK, the link between migration possibility educational choices is likely to be weak and that host countries clearly cream-skim the best applicants; neither of the two crucial premises is thus likely to hold. Bhargava, Docquier and Moullan (2011) find only a small positive effect of migration prospects on the decision to undertake medical training, clearly insufficient to generate a sizable effect on a county's stock of doctors. As noted in Docquier and Rapoport (2012), curtailing medical brain drain would, overall, increase staffing levels in developing countries.

[9] At the moment, simple estimates of the number of CHWs and their patients vary widely according to the source of information. Important efforts are nonetheless being undertaken in order to compile such data (for, for example by the One Million CHWs campaign: http://1millionhealthworkers.org).

[10] Budgetary constrains are associated with endemic shortages within medical and health science school faculties, lack of equipment and essential infrastructure maintenance. This, in turn, limits the number of health professionals, such as doctors and CHWs, trained in Sub-Saharan African countries.

further justified by official WHO recommendations, the bottom line conclusion of which is that "*Only broad and inclusive multi-sectoral planning at the national level will allow the coordination necessary to effectively scale up numbers and align health professional education with country health needs*", WHO (2011, p.18).

4.1 Model set-up

Consider a population that suffers from illnesses of varying severity, denoted by s, where $s \in [1, \infty)$. Let the health status of an individual with illness of severity s be $\frac{1}{s}$. Perfect health is the state valued at unity, where $s = 1$, and health status tends to zero as illness severity tends to infinity. Moreover, the distribution of the severity of illness in the population is given by a cumulative distribution function $F(s)$, defined on $[1, \infty)$.

We assume there are two kinds of health worker: doctors and Community Health Workers (CHWs). Doctors are indexed as type 1 and CHWs as type 2. A health worker of type *i*, where $i \in \{1, 2\}$, provides care of quality q_i, where $q_1 > q_2$. CHWs are capable of treating illnesses in the interval $[1, \hat{s}]$, while doctors are capable of treating all illnesses. If a health worker of type i expends time t treating a patient with illness of severity s, the health status of the treated patient will be:

$$\frac{(1 + q_i t)}{s}. \qquad (1)$$

Thus, to bring a patient to full health, when she is treated by a health worker of type i, requires a treatment time of:

$$t = \frac{s - 1}{q_i}. \qquad (2)$$

The time required is inversely proportional to the quality of the health worker. The cost of training a health care worker of type i is c_i. Define the quality-adjusted cost as $r_i = c_i / q_i$. We assume that:

$$r_1 > r_2. \qquad (3)$$

Inequality (2) is the usual assumption that producing a valued output (in this case, quality of care) comes at increasing marginal cost. We denote the budget available for training health care workers by $\bar{M}$, measured in country per capita terms. If m_i is the fraction of the population trained to be a health care worker of type i, then the budget constraint is:

$$c_1 m_1 + c_2 m_2 \leq \bar{M} . \tag{4}$$

m_i can be interpreted as the man-hours available to provide health care services of type i to the population. The total time spent on patients by health workers of type i must therefore not exceed m_i.

We measure social welfare (as far as health is concerned) as the mean of the logarithms of the health statuses of the population[11]. Define the *post-treatment health status* of an individual with illness of severity s as $h(s)$; of course, this depends upon the type of health worker assigned to him, and the time spent on treatment. The social objective is then:

$$W = \int_1^\infty \log h(s)\, dF(s) , \tag{5}$$

and the optimization problem is to decide how many doctors and CHWs to train, and how to assign them to treating patients with various severities of disease. Note that, if everyone in the population were brought to full health, then post-treatment health status would be $h(s) = 1$ for all, and the expression in (5) would be zero. Therefore, in general, the expression in (5) is negative, and we can therefore view $-W$ as the post-treatment *burden of disease* in the population, for this is the precisely the amount by which disease reduces the welfare of the population. Using this terminology, we can view

$$\int_1^\infty \log h(s)\, dF(s) - \int_1^\infty \log(1/s)\, dF(s) \tag{6}$$

as the amount by which the burden of disease in the society is reduced by health care.

A preliminary step for allocating resources optimally is the observation that:

[11] More on the choice of the objective below.

Lemma 1 *Optimal use of health care resources implies that CHWs treat only patients with illness severity in* $[1,\hat{s}]$ *and doctors treat only illnesses with severity in the interval* $(\hat{s},\infty)$.

The first part of Lemma 1 is true by definition, since CHWs are incapable of treating illnesses more severe than $\hat{s}$. It is the second part that is substantive; doctors will only treat more severe illness. Intuitively, the increasing marginal cost of quality implies it is always more cost-effective to allocate CHWs to less severe cases[12]. For the proof of Lemma 1 and subsequent theorems see the Appendix.

The complete resource allocation problem is to choose functions $t_1(\cdot),\ t_2(\cdot)$ and numbers $m_1,\ m_2$ to maximize[13]

$$\int_1^{\hat{s}} \log(\frac{1+q_2 t_2(s)}{s})\, dF(s) + \int_{\hat{s}}^{\infty} \log(\frac{1+q_1 t_1(s)}{s})\, dF(s)$$

subject to constraints (i)-(v), where Lagrangian multipliers are listed in parentheses:

$$\begin{array}{lll} \text{(i)} & \int_1^{\hat{s}} t_2(s)\, dF(s) \le m_2 & (\mu_2) \\ \text{(ii)} & \int_{\hat{s}}^{\infty} t_1(s)\, dF(s) \le m_1 & (\mu_1) \\ \text{(iii)} & 1+q_1 t_1(s) \le s, \quad s \in (\hat{s},\infty) & (\beta(s)) \\ \text{(iv)} & 1+q_2 t_2(s) \le s, \quad s \in [1,\hat{s}] & (\alpha(s)) \\ & t_1(s) \ge 0, \quad t_2(s) \ge 0 & \\ \text{(v)} & c_1 m_1 + c_2 m_2 \le \bar{M} & (\lambda) \end{array} \tag{7}$$

We define:

$$Q(x) = (1-F(x))x + \int_1^x s\, dF(s)\,, \tag{8}$$

[12] It is interesting that this result applies generally, for all budget levels and distributions of illness, hinging on the increasing marginal cost of quality of healthcare. This may explain why CHWs are trained and deployed also in rich countries, such as the USA.

[13] The objective function of program (7) can be written this way by Lemma 1.

where Q is a 'truncated mean' function. It is now possible to characterize the optimal resource allocation under certain premises. In particular:

Proposition 1 *Suppose that:*

(i) $r_1 > r_2$, *and*

(ii) $r_2 Q\left(\frac{r_1}{r_2}\right) - \left(r_2 F\left(\hat{s}\right) + r_1\left(1 - F\left(\hat{s}\right)\right)\right) < \bar{M} < r_2 Q\left(\hat{s}\right) - \left(r_2 F\left(\hat{s}\right) + r_1\left(1 - F\left(\hat{s}\right)\right)\right)$

Then the solution to program (7) *is given by:*

$$t_2(s) = \begin{cases} (s-1)/q_2, & s \in [1, s^*] \\ (s^*-1)/q_2, & s \in (s^*, \hat{s}] \end{cases}$$
$$t_1(s) = \frac{r_2 s^*}{c_1} - \frac{1}{q_1}, \quad s \in (\hat{s}, \infty) \tag{9}$$

where s^* *is the unique solution of the equation:*

$$\bar{M} = r_2 Q(s^*) - \left(r_2 F(\hat{s}) + r_1(1 - F(\hat{s})\right). \tag{10}$$

The values of (m_1, m_2) *are given by constraints* (i) *and* (ii) *in program* (7), *both of which are binding at the solution.*

Qualitatively, Proposition 1 states that, under premises (i)-(ii), the optimal solution has the following features: there is a severity s^* in the interval $[1, \hat{s}]$ such that CHWs spend sufficient time on those with illness severities $s \in [1, s^*]$ to completely cure them (i.e. raising their health status to 1); for any illness with severity in the interval $(s^*, \hat{s})$ CHWs spend a fixed, unvarying amount of time (and so these patients are not brought up to full health). Doctors treat all patients with illness more severe than $\hat{s}$, but again spend a constant amount of time on each case, and bring none of these patients up to full health. Note, in particular, that all patients receive treatment at the optimal solution.

The last sentence provides a justification for why we choose to maximize the average logarithm of post-treatment health status. If we had instead maximized the average post-treatment health status (not logged), then, it turns out, the optimal solution entails that patients at each illness severity *s* are either brought up to perfect health *or are not treated at all.* Because we never observe this kind of *bang-bang* solution, it is more realistic to apply a concave transformation to health status to form the health ministry's objective.

Proposition 1 presents the solution to program (7) when the parameters of the problem are in a set defined by premise (ii). As is usual for such problems, there will be different solutions to the program, depending upon precisely what the vector of parameters is. We do not attempt to provide a full characterization of the optimal solution for any possible parameter vector: our task here is not to advise fully the health ministry, but to show certain characteristics of the solution. Condition (ii) says that the budget $\bar{M}$ is not too large and not too small. It is difficult to understand why the condition takes this particular form without reading the proof.

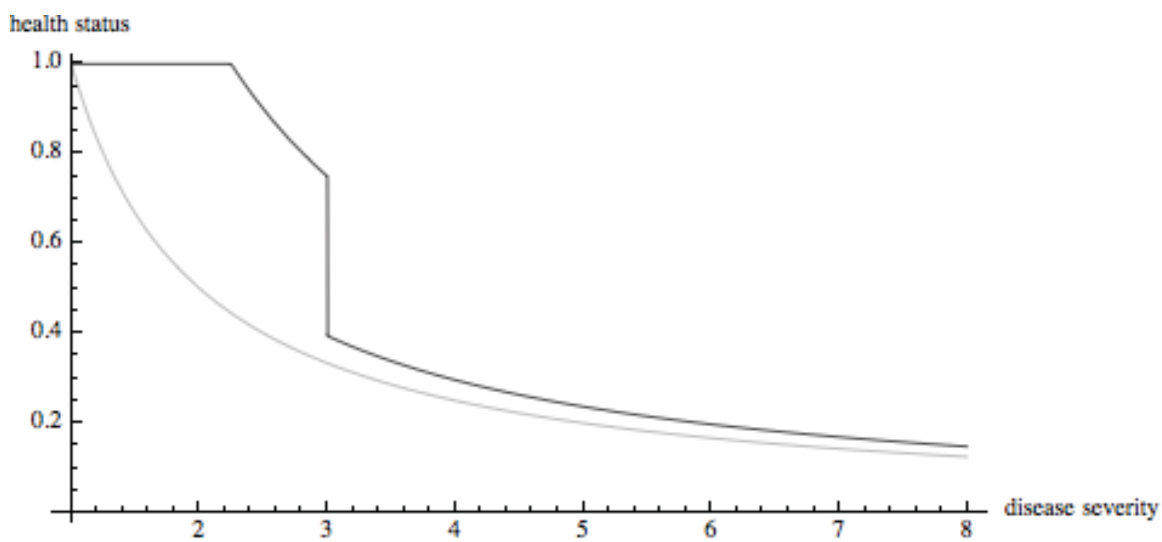
We consider an example, with $F(s) = 1 - \frac{1}{s}$ and $f(s) = \frac{1}{s^2}$, where f is the density of F. Set $(\hat{s}, c_1, q_1, c_2, q_2, r_1, r_2) = (3, 20, 10.526, 5, 5, 1.9, 1)$. Premise (ii) holds precisely when:

$$0.342 < \bar{M} < 0.799.$$

Set $\bar{M} = 0.51$. Then solving (10) gives $s^* = 2.248$. Note that $F(\hat{s}) = \frac{2}{3}$, so CHWs treat that fraction of the population. All patients with illnesses of severity less 2.248 are fully cured by treatment from CHWs. These comprise 55.5% of the population. The remaining patients with illness of greater severity are treated, but not restored to full health. The optimal supplies of the two kinds of medical personnel are $m_1 = 0.58\%$ and $m_2 = 7.9\%$, reported in population-percentage terms. In this example, there are many more CHWs than physicians.

We plot the pre-treatment and post-treatment health-status, as a function of *s*, for the numerical example at the optimal solution, in Figure 2. There are three regions: for $s < s^* = 2.248$, CHWs treat patients to full health; in the region $s^* < s < \hat{s} = 3$, CHWs spend the same amount of time on each patient, improving health status but not restoring patients to full health; finally, for $s > \hat{s}$, patients are treated by doctors, who spend an equal amount of time on each patient, again not bringing their patients up to full health. There is a saltus downward of health status at $\hat{s}$, from $(s^* / \hat{s})$ to $\frac{r_2}{r_1}(s^* / \hat{s})$.

Figure 2. Health status, before treatment (light curve) and after treatment (dark curve) at the optimal solution for numerical example



In general, social welfare at the optimal solution can be computed based on equation (5) and is given by:

$$W = \int_{1}^{s^*} \log(1)\, dF(s) + \int_{s^*}^{\hat{s}} \log\left(\frac{1+s^*-1}{s}\right) dF(s) + \int_{\hat{s}}^{\infty} \log\left(\frac{1+(r_2 s^*/r_1)-1}{s}\right) dF(s) = \int_{s^*}^{\hat{s}} \log(s^*/s)\, dF(s) + \int_{\hat{s}}^{\infty} \log\left(\frac{r_2 s^*}{r_1 s}\right) dF(s). \quad (11)$$

It is interesting that health status takes a saltus downward at $\hat{s} = 3$. So patients whose illness is slightly less severe than $\hat{s}$ are 'significantly' better off, treated by the CHW, than patients whose illness is slightly more severe than $\hat{s}$, who are treated by the doctor. This will not be the case for *all* possible problems. Proposition 1 only characterizes the optimal solution for problems characterized by premise (ii).

4.2 Medical migration

This section examines comparative statics of the optimal solution characterized in Proposition 1 in the context of outward medical migration of doctors, who are trained at public expense, but then take jobs in rich countries. In contrast, CHWs are assumed to remain geographically immobile. These assumptions are backed by the strong evidence presented in section 1, which points to high rates of outward migration of fully qualified physicians from developing countries, whereas CHWs lack the formal qualifications to fill medical posts abroad.

Denote by π the fraction of doctors who, after training, stay in the country, and by $1-\pi$ the fraction that migrates. The effect of migration in optimization problem (7) is simply to change the effective cost of training a doctor from c_1 to c_1/π. The effective cost of producing m_1 doctors who stay in the country is thus $c_1 m_1/\pi$ (i.e. if one trains m_1/π doctors, the number who stay will be $\pi(m_1/\pi) = m_1$).

4.2.1 Resource allocation with migration

In the case of Proposition 1, the equations that characterize the optimal values $\left(s^*, m_1, m_2\right)$ are:

$$\int_1^{s^*} \frac{s-1}{q_2} dF(s) + \int_{s^*}^{\hat{s}} \frac{s^*-1}{q_2} dF(s) - m_2 = 0 \qquad (12)$$

$$((\pi r_2 s^* / c_1) - (1/q_1))(1 - F(\hat{s})) - m_1 = 0 \qquad (13)$$

$$r_2 Q(s^*) - (r_2 F(\hat{s}) + \frac{r_1}{\pi}(1 - F(\hat{s})) - \bar{M} = 0, \qquad (14)$$

which comprise three equations in the three unknowns $\left(s^*, m_1, m_2\right)$.

Equation (12) is the binding constraint (i) of program (7) at the optimal solution, (13) is binding constraint (ii) of (7), and (14) is equation (10). Throughout, c_1/π has been substituted for c_1.

The Jacobian of this system with respect to these three variables is:

$$J = \begin{pmatrix} \dfrac{F(\hat{s}) - F(s^*)}{q_2} & 0 & -1 \\ \pi r_2 (1 - F(\hat{s}))/c_1 & -1 & 0 \\ r_2(1 - F(s^*)) & 0 & 0 \end{pmatrix}.$$

We now differentiate the three equations with respect to π, which gives the vector:

$$b = \begin{pmatrix} 0 \\ r_2 s^* (1 - F(\hat{s}))/c_1 \\ \dfrac{r_1(1 - F(\hat{s}))}{\pi^2} \end{pmatrix}.$$

By the implicit function theorem, the derivatives of the optimal values of $\left(s^*, m_1, m_2\right)$ with respect to π are given by:

$$\begin{pmatrix} \partial s^* / \partial\pi \\ \partial m_1 / \partial\pi \\ \partial m_1 / \partial\pi \end{pmatrix} = -J^{-1}b, \tag{15}$$

or $J\begin{pmatrix} \partial s^* / \partial\pi \\ \partial m_1 / \partial\pi \\ \partial m_1 / \partial\pi \end{pmatrix} = -b$. Solving these three equations for the derivatives at $\pi = 1$ gives:

$$\begin{aligned} \left.\frac{\partial s^*}{\partial\pi}\right|_{\pi=1} &= -\frac{r_1(1-F(\hat{s}))}{r_2(1-F(s^*))} < 0 \\ \left.\frac{\partial m_1}{\partial\pi}\right|_{\pi=1} &= -\frac{(1-F(\hat{s}))^2}{q_1(1-F(s^*))} + \frac{r_2 s^*(1-F(\hat{s}))}{c_1} > 0 \\ \left.\frac{\partial m_2}{\partial\pi}\right|_{\pi=1} &= -\frac{r_1(1-F(\hat{s}))(F(\hat{s})-F(s^*))}{c_2(1-F(s^*))} < 0 \end{aligned} \tag{16}$$

This means that as π *decreases* from a value of one, s^* *increases*, and *more* CHWs are trained. Because m_2 increases as π decreases, it immediately follows from the budget constraint that $\frac{\partial m_1}{\partial\pi} > 0$, a fact that can also be calculated from (16). In sum, we have shown:

Proposition 2. *Under the premises of Proposition* 1, *as the fraction of migrating doctors increases from zero, then*:

(a) *the number of* CHWs *trained increases*,

(b) *the number of doctors trained decreases*,

(c) *those will illness severity* $s \leq \hat{s}$ *receive more treatment in aggregate, and*

(d) *those will illness severity* $s > \hat{s}$ *receive less treatment*.

The effect on the burden of disease from migration of doctors is, of course, of key interest. We thus compute the effect on burden from migration. From equation (11), when migration is included, we can write welfare at the optimal solution as:

$$W(\pi) = \int_{s^*}^{\hat{s}} \log \frac{s^*(\pi)}{s} dF(s) + \int_{\hat{s}}^{\infty} \log \frac{r_2 s^*(\pi)\pi}{r_1 s} dF(s)$$

and so we compute the derivative as:

$$\left.\frac{dW(\pi)}{d\pi}\right|_{\pi=1} = \frac{ds^*}{d\pi}\left(\frac{1-F(s^*)}{s^*}\right) + 1 - F(\hat{s}) = (1-F(\hat{s}))(1-\frac{r_1}{r_2 s^*}) > 0, \qquad (17)$$

where we used the expression in (16) for $\frac{ds^*}{d\pi}$ and concluded the last inequality from the fact that $s^* > \frac{r_1}{r_2}$. Expressing this change in welfare as an elasticity, we have:

$$\left.\frac{dW}{d\pi}\frac{\pi}{|W|}\right|_{\pi=1} = (1-F(\hat{s}))(1-\frac{r_1}{r_2 s^*}) / \left| \int_{s^*}^{\hat{s}} \log(s^*/s)\, dF(s) + \int_{\hat{s}}^{\infty} \log\left(\frac{r_2 s^*}{r_1 s}\right) dF(s) \right|. \qquad (18)$$

This can be illustrated using the example described earlier. Evaluating this elasticity in that context gives:

$$\left.\frac{dW}{d\pi}\frac{\pi}{|W|}\right|_{\pi=1} = 0.078\,.$$

This means that if π falls from one to 0.9, welfare will decrease (or the burden of disease will increase) by approximately 0.78%.

In general, the effect of an increase in physician migration on the number of doctors and CHWs at the optimal solution for the country can be found by computing the elasticities of m_1 and m_2 with respect to π. These are evaluated using the equations in (16):

$$\left.\frac{dm_1}{d\pi}\frac{\pi}{m_1}\right|_{\pi=1} = -\frac{(1-F(\hat{s}))^2}{m_1 q_1 (1-F(s^*))} + \frac{r_2 s^*(1-F(\hat{s}))}{c_1 m_1}, \qquad (19)$$

and

$$\left.\frac{dm_2}{d\pi}\frac{\pi}{m_2}\right|_{\pi=1} = -\frac{r_1(1-F(\hat{s}))(F(\hat{s})-F(s^*))}{c_2 m_2(1-F(s^*))} \tag{20}$$

As an illustration, in terms of our example, these elasticities evaluate to:

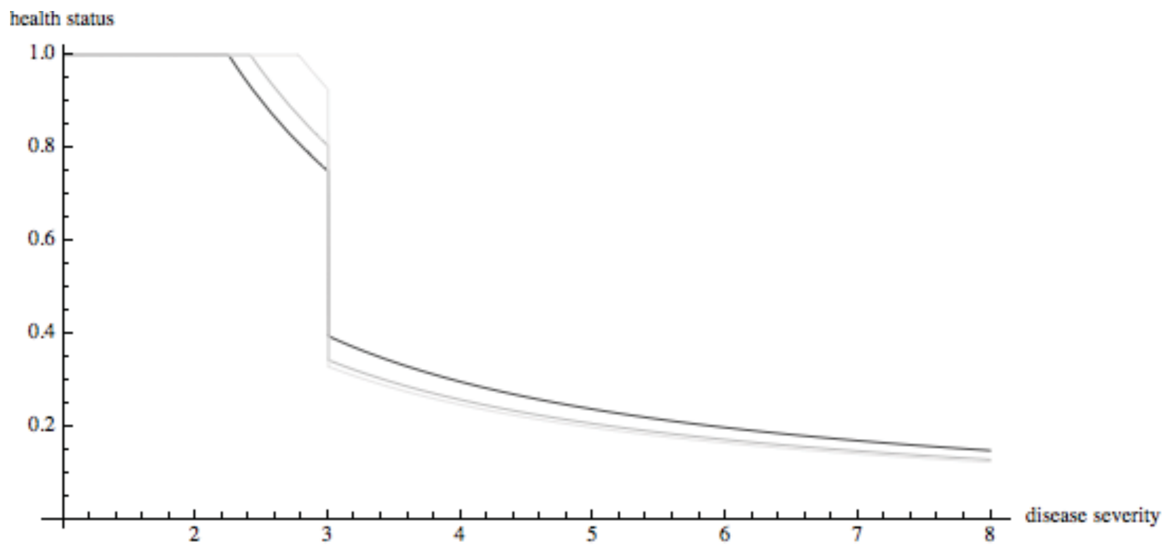
$$\left.\frac{dm_1}{d\pi}\frac{\pi}{m_1}\right|_{\pi=1} = 2.37, \quad \left.\frac{dm_2}{d\pi}\frac{\pi}{m_2}\right|_{\pi=1} = -0.403. \tag{21}$$

In other words, a fall in the fraction of doctors staying in the country from unity to 0.9 reduces the number of doctors *practicing* in the country by 23.7 %, at the optimal allocation, and *increases* the number of CHWs trained by 4.0%. Because of the effective increase in the price of training doctors, the fraction of doctors trained falls (in this example) by more than twice the loss due to migration. Moreover, from the first derivative in (16), the effect of migration is to *increase* s^*. The increase in the number of CHWs implies that patients with illnesses in the interval $[0,\hat{s}]$ actually *are better off* with some physician migration; they have more CHWs to treat them. The brunt of the increase in the burden of disease is borne entirely by patients with illnesses too severe for CHWs to treat. This result is comprehensible if we recall that the objective of program (7) is to maximize an average of 'utilities' in the population. As doctors become effectively more expensive due to the leakage of migration, it is optimal to substitute CHWs for doctors: but since CHWs are constrained to treat only relatively minor illnesses, patients with those illnesses have improved outcomes.

4.2.2 Health outcomes with migration

Let us return to our example. Now we incorporate outward migration by assuming $\pi = 0.9$ and $\pi = 0.75$; that is, 10% or 25% of the doctors migrate. Solving for the new severity threshold and optimal resource allocation between training doctors and CHWs allows us to graph the post-treatment health status distribution these two levels of migration. We graph the post-treatment health status of the population optimal solutions with and without migration in Figure 3.

Figure 3. Health status at zero migration (dark curve) , 10% migration (gray curve) and 25% migration (lightest curve) as a function of *s* at the optimal solutions



The effect of migration on the post-treatment health distribution summarized in Proposition 2 can be observed in the figure. Patients who are treated by CHWs actually *do better* with migration – that is, more of them are restored to full health. Patients whose severity is greater than $\hat{s}$ do worse. In other words, migration induces the Ministry of Health to cut back on doctor training, and shift resources towards training more CHWs. The threshold s^* increases with migration, as we saw in (16). The overall burden of disease, of course, increases – for this example from 0.6588 to 0.6637 at 10% migration, by 0.74%, which agrees well with the estimate based on the elasticity of welfare at $\pi = 1$, given above, of 0.78%.

In general, the basic intuition behind these patterns, driven by (16), is as follows. Increasing migration is reflected in an increase in the effective cost of training doctors – for the cost of training a doctor who will be available in the country is $\frac{c_1}{\pi}$. Consequently, an increase in migration increases the relative cost of the more expensive healthcare input, and so the Ministry economizes by hiring fewer of them. In fact, after migration, in the optimal solution, the Ministry spends less in total on training doctors than before, implying more expenditure on training CHWs given an unchanged budget. This is why those who are not severely ill actually benefit from the migration of doctors.

4.2.3 Resource allocation and health outcomes with migration and reimbursement

Now suppose migrating doctors are taxed to pay back the cost of their training, or the donor country receives a reimbursement from the countries to which they migrate, equal to the cost of their training. Migration thus increases the budget available to the Ministry to train doctors and CHWs, giving rise to a new optimal resource allocation.

Let n_1 be the number of doctors who are trained. The number who stay to practice in the home country is $m_1 = \pi n_1$. The cost of training the physicians who migrate is $c_1(1-\pi)n_1$. We now assume that this amount is reimbursed to the Ministry of Health by the rich countries to which their doctors have migrated. Therefore the budget constraint for the Ministry becomes:

$$c_1 n_1 + c_2 m_2 = \bar{M} + c_1(1-\pi)n_1 \text{ or } c_1(\pi n_1) + c_2 m_2 = \bar{M} \quad . \qquad (22)$$

The new optimization program for the Ministry is exactly the same as program (7) except that the budget constraint (v) is replaced with equation (22) and constraint (ii) is replaced with:

$$\int_{\hat{s}}^{\infty} t_1(s)\,dF(s) \le \pi n_1. \qquad (23)$$

But this new program is identical to program (7) except that m_1 has been replaced with πn_1. Therefore, the solution to program (7) (without migration) will be identical to the solution of the new program, with $\pi n_1 = m_1^*$, where m_1^* is the optimal supply of physicians in program (7), absent migration. Thus, if the cost of training the migrants is reimbursed to the home country, then there will be no change in medical care or the burden of disease.

However, there is an important *ceteris paribus* assumption hidden here – that increasing the number of physicians that the home country trains will not decrease the quality of trained physicians. More medical schools will be needed, the applicant pool will be larger, etc., so in all likelihood the *ceteris paribus* assumption is false. This means that, in reality, the home country should be reimbursed more than the cost of training the migrating physicians, if its burden of disease is not to increase.

It is clear that if the home country is reimbursed *less* than the cost of training the migrants, then the burden of disease increases there, for this would be equivalent to decreasing the medical budget from the case of full reimbursement, which therefore must decrease the value of optimization program.

5. Conclusion

A shortage of medical personnel has been addressed in developing countries through the systematic training and deployment of CHWs who supplement healthcare provision by fully qualified doctors. Our analysis develops a model of a two-tiered structure of healthcare provision and characterizes the optimal allocation of resources between training doctors and CHWs, as well as the implications for population health outcomes. Outward medical migration of physicians distorts the cost of training doctors relative to geographically immobile CHWs, thereby shifting resources towards training CHWs. Since CHWs can only treat a limited range of illnesses, the additional investment in training of CHWs can only give rise to additional treatment of relatively low severity illness.

While migration increases the burden of disease in society overall, it stimulates health care provision at the lower end of the illness-severity spectrum, improving health outcomes for those patients; sufferers of relatively severe medical conditions who can only be treated by doctors are made worse off[14]. This provides insight on an important policy debate, centered on whether foreign aid should be used in system-wide interventions, aimed at strengthening the entire health system, or, rather, on disease-specific programs, aimed at particular health conditions, such as AIDS or malaria (see Warren et al. 2013). Our results show that patients affected by diseases that require the attention of fully qualified doctors are particularly harmed by medical emigration. This provides a novel justification for disease-specific interventions.

We show that, under a *ceteris paribus* assumption, if the donor country is reimbursed by the training cost of emigrating physicians, the overall burden of disease in society is maintained at its pre-migration level. Because the *ceteris paribus* assumption does not take into account the increasing marginal costs of expanding medical schools, and the decrease in the quality of medical students if the applicant pool is enlarged, in reality the donor country must be reimbursed by more than the cost of training its migrating physicians to remain whole.

In a hypothetical world where donor countries could prevent migration, the number of emigrating doctors and the level of reimbursement would be the outcome of a bargaining game between donor and host. In such a setting, the donor country would be reimbursed by more than is required to restore overall population health. This provides a normative justification for the view that recipients of fully qualified medical personnel through migration should substantially compensate developing countries.

In the previous sections we present this reimbursement as lump sum transfer. This need not be the case, as demonstrated by recent doctors retention policies, such as Malawi's Emergency Human Resources Programme (EHRP). As shown in Table 2, the density of physicians in this country is low and, simultaneously, a large share of the doctors trained in

[14] Several middle income countries achieved important increases in life expectancy through cost-effective primary care provision, relying heavily on CHWs. However, this cost-effective strategy is insufficient for tackling more severe non-communicable diseases that typically require treatment by fully qualified doctors, as made clear in a recent OECD policy report about China (OECD, 2013).

Malawi emigrate. In order to curtail this deficit of medical personnel, international partners, namely the UK's Department for International Development (DFID) and the Global Fund to Fight AIDS, Malaria and Tuberculosis, supported Malawi by providing aid (over 100 million dollars between 2004 and 2009) used to subsidize retention policies for doctors (DFID, 2010). First, EHDR provided a substantial top-up of salaried medical doctors in the country (roughly 50% of their before salary), in order to reduce the incentives for migration. Second, EDHR funded the training of medical doctors; this policy is associated[15] with a large increase in the number of medical doctors on training in the country[16].

References

Ashraf, N., Bandiera O. and Lee, S. 2013. Selection and Compensation of Community Health Workers in Zambia. JPAL working paper.

Astor, A., Akhtar, T.,Matallana, M. A.,Muthuswamy, M., Olowu, F. A., Tallo, V. 2005. Physician migration: views from professionals in Colombia, Nigeria, India, Pakistan and the Philippines. Social Science & Medicine, 61, 2492–2500

Acemoglu, D. and Johnson, S. Disease and development. 2007. Journal of Political Economy 115, December 2007: pp. 925-985.

Banerjee, A., Deaton, A. and Duflo, E. 2004,. Wealth, health, and health services in rural Rajasthan. American Economic Review, papers and proceedings, 94(2), 326-330.

Beine M., Docquier, F. and Rapoport H. 2001. Brain drain and economic growth: theory and evidence. Journal of Development Economics, 64: 275-89.

Bhagwati, J. and Hamada, K. 1974. The brain drain, international integration of

[15] Policy reports (DFID, 2010) claim that EDHR caused the number of medical doctors practicing in the country to increase noticeably; nonetheless, a causal impact evaluation of this program has not been undertaken.

[16] It could be argued that another form of compensation might consist of the host countries supplying doctors do developing countries. However, with some exceptions (such as the Cuban Medical Cops) the deployment of medics to developing countries is often transitory and associated with calamity relief and disease specific programs.

markets for professionals and unemployment, Journal of Development Economics, 19-42.

Bhargava, A. and Docquier, F. 2008. HIV pandemic, medical brain drain, and economic development in Sub-Saharan Africa. World Bank Economic Review, 22: 345-66.

Bhargava, A., Docquier, F. and Moullan, Y. 2011. Modeling the effect of physician emigration on human development, Economics and Human Biology, 9(2):172-83.

Bundred, P. and Levitt, C. 2000. Medical migration: who are the real losers?, The Lancet 356, 9225: 245-46.

Cutler, D., Deaton, A. and Lleras-Muney, A. 2006. The Determinants of Mortality. Journal of Economic Perspectives, 20(3): 97-120.

Deaton, A. 2011. What does the empirical evidence tell us about the injustice of health inequalities? Working Papers 1364, Princeton University, Woodrow Wilson School of Public and International Affairs, Center for Health and Wellbeing.

DFID. 2010. Evaluation of Malawi's Emergency Human Resources Programme.

Deaton, A. 2013. The Great Escape: Health, Wealth, and the Origins of Inequality. Princeton University Press.

Docquier, F. and Rapoport, H. 2012. Globalization, Brain Drain, and Development. Journal of Economic Literature, 50(3): 681-730.

Kangasniemi, M., Winters, L.A. and Commander, S. 2007. Is the medical brain drain beneficial? Evidence from overseas doctors in the UK. Social Science & Medicine 65 (2007) 915–923.

Kinfu, Y., Dal Poz, M., Mercer, H. and Evans, D. 2009. The health worker shortage in Africa: are enough physicians and nurses being trained? Bulletin of the World Health Organization (87): 225-230.

McCulloch, R. and Yellen, J. 1975. Consequences of a tax on the brain drain for unemployment and income inequality in less developed countries. Journal of Development Economics, 249-64.

McCulloch, R. and Yellen, J. 1977. Factor mobility, regional development and the distribution of income, Journal of Political Economy, 85, 1: 79-96.

McHale, J. 2009. Taxation and skilled Indian migration to the United States: revisiting the Bhagwati tax, Chapter 12 in J. Bhagwati and G. Hanson (eds): Skilled immigration: problems, prospects and policies, Oxford University Press, 362-86.

Mills, E.J., Kanters, S., Hagopian, A., Bansback, N. Nachega, J., Alberton, M., Au-Yeung, C. Mtambo, A., Bourgeault,I., Luboga, S., Hogg, R., Ford, N. 2011. The financial cost of doctors emigrating from sub-Saharan Africa: human capital analysis. British Medical Journal, 343:d7031.

Mountford, A. 1997. Can a brain drain be good for growth in the source economy? Journal of Development Economics, 53(2), 287–303.

Mullan, F., Frehywot, S., Omaswa, F., Buch, E., Chen, C., Greysen, R., Wassermann, T., ElDin ElGaili Abubakr, D., Awases, M., Boelen, C., Diomande, M., Dovlo, D., Ferro, J., Haileamlak, A., Iputo, J., Jacobs, M., Koumaré, A.K., Mipando, M., Monekosso, G., Olapade-Olaopa, E., Rugarabamu, P., Sewankambo, N., Ross, H., Ayas, H., Chale, S., Cyprien, S., Cohen, J., Haile-Mariam, T., Hamburger, E., Jolley, L., Kolars, J.C., Kombe, G., Neusy, A.J. 2011. Medical schools in sub-Saharan Africa. 377: 1113–21.

World Health Organization .1989. Strengthening the performance of community health workers in primary health care. Report of a WHO Study Group. Geneva, World Health Organization, WHO Technical Report Series, No. 780.

World Health Organization – Global Health Workforce Alliance. 2010. Malawi's Emergency Human Resources Programme, Geneva,

World Health Organization. 2011. Transformative scale-up of health professional education - An effort to increase the numbers of health professionals and to strengthen their impact on population health.

Rodriguez, C. 1975. Brain Drain and Economic Growth: A dynamic model.
Journal of Development Economics, 223-248.

Vidal, J. 1998. The effect of emigration on human capital formation. Journal of Population Economics, 11: 589-600.

Warren, A. Wyss, K., Shakarishvili, G., Atun, R. and De Savigny, D. 2013. Global health initiative investments and health systems strengthening: a content analysis of global fund investments, *Global Health* 9:30

World Health Organization. 2006. The World Health Report: working together for health.

World Health Organization. 2010. Global experience of community health workers for delivery of health related Millennium development goals: a systematic review, country case studies, and recommendations for integration into national health systems

Appendix

Proofs of theorems are provided below.

Proof of Lemma 1:

Suppose doctors spend time $t(s)$ treating patients with illnesses in the interval $s \in (s_1, s_1+\delta) \subset [1,\hat{s}]$. The outcome for this group of patients is

$$\int_{s_1}^{s_1+\delta} \log\left(\frac{1+q_1 t(s)}{s}\right) dF(s)\,. \tag{A1}$$

The cost of this treatment is $c_1 \int_{s_1}^{s_1+\delta} t(s)\,dF(s) = C_1$. Now let $\hat{t}(s) = \frac{q_1 t(s)}{q_2}$, and let these patients be treated instead by CHWs with treatment times $\hat{t}(s)$. The welfare outcome is identical but the cost of treatment is:

$$c_2 \int_{s_1}^{s_1+\delta} \frac{q_1 t(s)}{q_2} dF(s) = \frac{c_2 q_1}{q_2}\frac{C_1}{c_1} = \frac{r_2}{r_1} C_1 < C_1\,, \tag{A2}$$

where the inequality follows from assumption (3). Hence it is not optimal to treat these patients using doctors. ■

Proof of Proposition 1:

1. Observe first that $Q'(x) = 1 - F(x) > 0$, so Q is increasing.

2. The proof is based on the fact that program (7) is a convex program. Define the Lagrangian function:

$$\begin{aligned}
L(\varepsilon) = {} & \int_1^{\hat{s}} \log\left(\frac{1+q_2(t_2(s)+\varepsilon\Delta t_2(s))}{s}\right) dF(s) + \int_{\hat{s}}^{\infty} \log\left(\frac{1+q_1(t_1(s)+\varepsilon\Delta t_1(s))}{s}\right) dF(s) + \\
& \mu_2\left(m_2 + \varepsilon\Delta m_2 - \int_1^{\hat{s}} (t_2(s)+\varepsilon\Delta t_2(s))\,dF(s)\right) + \mu_1\left(m_1 + \varepsilon\Delta m_1 - \int_1^{\hat{s}} (t_1(s)+\varepsilon\Delta t_1(s))\,dF(s)\right) + \\
& \int_1^{s^*} \alpha(s)\big(s - (1+q_2(t_2(s)+\varepsilon\Delta t_2(s)))\big) dF(s) + \lambda\left(\bar{M} - c_1(m_1+\varepsilon\Delta m_1) - c_2(m_2+\varepsilon\Delta m_2)\right)
\end{aligned}$$

where the functions $t_1(s), t_2(s)$ and the numbers m_1 and m_2 comprise the candidate for the optimal solution of the program. If we can produce a non-negative function $\alpha(\cdot)$ on $[1, s^*]$

and non-negative constants (μ_1,μ_2,λ) such that $L'(0)=0$ at the values of these variables stated in the proposition, then the proposition is proved. For this will mean that the concave function L is maximized at $\varepsilon=0$. In the Lagrangian function L the functions $\Delta t_i(\cdot), i=1,2,$ are arbitrary feasible variations from the conjectured optimal solution, as are the numbers Δm_i.

2. Evaluating the derivative at zero:

$$L'(0)=\int_1^{\hat{s}}\frac{q_2\Delta t_2(s)}{1+q_2t_2(s)}dF(s)+\int_{\hat{s}}^{\infty}\frac{q_1\Delta t_1(s)}{1+q_1t_1(s)}dF(s)+\mu_2(\Delta m_2-\int_1^{\hat{s}}\Delta t_2(s)\,dF(s))+$$
$$\mu_1(\Delta m_1-\int_{\hat{s}}^{\infty}\Delta t_1(s)\,dF(s))-\lambda(c_1\Delta m_1+c_2\Delta m_2)-\int_1^{s^*}\alpha(s)q_2\Delta t_2(s)\,dF(s). \tag{A3}$$

If we can choose non-negative Lagrangian multipliers (μ_1,μ_2,λ) and a non-negative function $\alpha(s)$ on $[1,s^*]$ so that the coefficients of $(\Delta t_1(s),\Delta t_2(s),\Delta m_1,\Delta m_2)$ are all annihilated, then the result is proved.

3. The coefficients of these variations are:

$\Delta t_2(s):\quad \dfrac{q_2}{1+q_2t_2(s)}-\alpha(s)q_2\delta[1,s^*]-\mu_2=0,$ for $s\in[1,\hat{s}]$. $\delta[\mathrm{a,b}]$ is the function that is 1 on $[\mathrm{a,b}]$, and 0 elsewhere;

$\Delta t_1(s):\quad \dfrac{q_1}{1+q_1t_1(s)}-\mu_1=0$

$\Delta m_1:\quad \mu_1-c_1\lambda=0$

$\Delta m_2:\quad \mu_2-c_2\lambda=0$

4. From the Δt_2 condition, we have $\alpha(s)=-\dfrac{\mu_2}{q_2}+\dfrac{1}{s}$ for $s\in[1,s^*]$. At s^*, we choose $\alpha(s^*)=0$, implying $\mu_2=\dfrac{q_2}{s*}$. It follows that $\alpha(s)>0$ for $s<s^*$. Therefore

$\lambda=\dfrac{q_2}{c_2s^*}=\dfrac{1}{r_2s^*}$ and so $\mu_1=\dfrac{c_1}{r_2s^*}$. Therefore $(1+q_1t_1)\dfrac{c_1}{r_2s^*}=q_1$, giving $t_1=(\dfrac{r_2s^*}{r_1}-1)/q_1$.

For t_1 to be positive we require $s^*>\dfrac{r_1}{r_2}$, which we will attend to below. We must also check that we are not treating patients with illness severityt $s > \hat{s}$ with more treatment than they require: that is, we want $1 + q_1 t_1 \le \hat{s}$. But this is true because $s^* < \hat{s} < \frac{r_1}{r_2}\hat{s}$.

5. We now check the budget constraint for the proposed solution, which is:

$$c_2 \int_1^{s^*} \frac{s-1}{q_2} dF(s) + c_2 \int_{s^*}^{\hat{s}} \frac{s^* - 1}{q_2} dF(s) + c_1 \int_{\hat{s}}^{\infty} \left((\frac{r_2 s^*}{r_1} - 1) / q_1 \right) dF(s) = \bar{M} \ .$$

This can be written as :

$$r_2 \int_1^{s^*} (s-1) dF(s) + r_2 (s^* - 1)(F(\hat{s}) - F(s^*)) + (r_2 s^* - r_1)(1 - F(\hat{s})) = M$$

which in turn is equivalent to condition (10) defining s^* . Finally, our premise (ii) is exactly the condition that tells us a unique solution s^* exists to equation such that $\frac{r_1}{r_2} < s^* < \hat{s}$.

This is so because the function Q is monotone increasing, and the existence of s^* therefore follows from condition (ii) by the intermediate value theorem. ■