

Managing the supply of physicians' services through intelligent incentives

Brian R. Golden PhD, Rosemary Hannam MBA, Douglas Hyatt PhD

Addressing issues of access to health care for Canadians has been an ongoing challenge for policy-makers. Although Canada has experienced times of physician oversupply (e.g., in the 1990s),¹ which it could experience again as a result of delayed retirements, increased enrollment in medical schools and the arrival of international medical graduates,² uniformly adequate access to care remains a challenge. The state of access to health care has recently been reported by the Health Council of Canada's 2010 comparison of Canada with other developed countries in terms of access to after-hours medical care, wait times and the ability to get an appointment with a physician the same or next day.³

Central to the issue of access is the adequacy of the supply of physicians — specifically, whether the number of physicians and their work effort sufficiently addresses the health care needs of the population. Supply is appropriately managed when there is neither a shortage nor surplus of services. Governments, when attempting to address shortages, recognize that training and attracting new physicians is both time-consuming and expensive. Similarly, there is little recourse for immediately reducing the number of physicians when supply is too great. Governments, however, can have a more immediate impact on the supply of physicians by influencing physicians' decisions to adjust their workloads to either increase or reduce the number of services they provide.

Financial incentives can have profound effects on the behaviour of people supplying, or considering supplying, their services. However, recent reviews have concluded that the behavioural impact of such incentives is more complex than may sometimes be recognized.⁴⁻⁶ It is possible that the same compensation policies aimed at attracting workers and increasing the supply of labour may also induce workers to reduce their workloads, thereby reducing supply. There is a very real risk that paying more — with the intention of buying more — could be incentive for physicians to provide fewer services. Indeed, a recent paper reports that physicians' incomes in Canada are high and growing,⁷ yet there continue to be shortages of physicians throughout parts of the country.

The purpose of our paper is to introduce noneconomists to the concepts of labour supply theory in the context of physician remuneration, and to explain some of the fundamental principles that policy-makers can readily apply. We present a framework for understanding how financial incentives can influence the supply of physician services and show how it can be used to project reactions by the workforce to policy initiatives commonly used to influence supply. We do not address the frequently studied second-order question of whether physicians are providing appropriate care as a result of financial incentives; instead, we address the more fundamental first-order question of whether physicians offer their services and to what degree.

Framework for understanding physician supply

Although the focus of many of the policies to address perceived physician shortages has been increasing the number of physicians, Jeon and Hurley^{8,9} and Crossley and colleagues¹⁰ have persuasively argued that a wide array of additional factors affect supply, including opportunities for

Competing interests: Brian Golden has taught seminars on change management for the Canadian Medical Association's Physician Management Institute. No other competing interests were declared.

This article has been peer reviewed.

Correspondence to: Rosemary Hannam, rosemary.hannam@rotman.utoronto.ca

CMAJ 2012, DOI:10.1503/cmaj.111018

KEY POINTS

- Canadian governments use various incentives, including increasing wages or fees and providing nonwage benefits, to ensure sufficient supply of physicians.
- An increase in the wage rate induces both an income effect (which induces less work) and a substitution effect (which induces more work); it is not possible to predict which of these effects will dominate.
- Some physicians may initially increase their workload in response to a fee increase (substitution effect), but may reduce their workloads later in their career or retire from practice earlier (income effect).
- Unlike wage increases, nonwage benefits can be designed so that either a substitution or income effect dominates, depending on the government's objectives.

This paper is based on research surrounding the most recent Physician Service Agreements between the Ontario Ministry of Health and Long-Term Care and the Ontario Medical Association. The authors were asked by both parties to act as a neutral body providing advice, based on economic theory and focus group research, on the most effective ways to match the demand for physician services with supply.

practice versus nonpractice income, technology, attitudes toward work and the policy environment. Importantly, a distinction is made between the extensive margin (the supply of physicians) and the intensive margin (the services supplied by each physician); it is reductions in the latter that have been an important driver of the difficulties in accessing care. Jeon and Hurley conclude that

Although increases in the supply of providers may be important long-run policies, in the near term the most effective means of increasing access would be to increase hours of work and service provision by the existing stock of practicing physicians.¹¹

The supply of physician services at a point in time is determined by the number of physicians available to work, the number of hours devoted to providing services and the associated number of services provided per hour.

There are several policy options at the disposal of provincial and territorial governments. As has been seen previously in Canada, governments can attempt to influence attraction and retention by increasing fee schedule amounts to compete with those of other provinces, or by providing nonwage incentives (e.g., recruitment bonuses, relocation assistance, reimbursement of tuition and incentives targeting recent graduates).¹² In addition, these incentives may be used to affect where in the province a physician would practise, such as in underserved rural areas. Furthermore, the number of physicians practising in a province or territory can be influenced by changing the number of positions available in medical schools and residency programs and the process for licensing graduates of international medical schools.

To address work intensity, the government can change the rate and design of physician compensation and the services that are required in return. Aside from negotiating changes to fee schedules, governments have introduced alternative models of reimbursement (e.g., salary, capitation). The Canadian Institute for Health Information estimates that 39.1% of Canadian physicians are currently paid, at least partially, through alternate fee arrangements (up from 28.1% in 2000–2001).¹³ The impact of these changes will be influenced by the relative strengths of their income and substitution effects.

Income and substitution effects

The core of the theory of labour supply is that workers make choices between time spent working and time spent engaged in nonwork activities (leisure). Influencing these choices are preferences for income and leisure, wages and nonwage income. Although wage and nonwage income are only two factors that could influence the choice between time spent at work and time spent at

leisure, they are important because they can be directly influenced by policy-makers. However, the influence of changes in income and wages on behaviour is complex and sometimes ambiguous.

Wage increases may create an incentive for workers to either work more hours or work fewer hours, depending on their individual preferences, because an increase in wages has both income and substitution effects. These effects work in opposite directions, and the ultimate impact on the behaviour of workers will depend on which is the dominating effect. This basic framework has been extended by researchers, such as McGuire and Pauly,¹⁴ to consider the labour supply of physicians. The opposing income and substitution effects in response to changes in the wage rate are the basis for the backward-bending supply curve of labour. This theory makes four assumptions: workers choose their hours, workers are homogeneous, there are no contractual obligations, and workers are utility-maximizing agents. Briefly described, the theory states that when the substitution effect outweighs the income effect, wage increases up to a certain threshold will elicit an increase in the supply of labour; further increases beyond the threshold will lead to a reduction in the supply of labour.¹⁵

Consider first the income effect of a wage increase. An increase in wage translates to an increase in income. Assuming that time spent in nonlabour activities such as household work and leisure is valued by workers, and that they prefer to spend more time engaged in these other activities as their income increases, the income effect of a wage increase would induce a worker to work less.

The same wage increase could also engender a substitution effect. As wages increase, the amount of money a worker foregoes by not working increases; it implicitly becomes more expensive to reduce work time. Under the reasonable assumption that a worker would purchase less of an expensive good or service when its price increases, workers will substitute time in the labour market for time out of the labour market and work more hours.

Thus, an increase in wages induces an income effect, which induces a worker to work less, and a substitution effect, which induces a worker to work more. Which of these two opposing effects will dominate ultimately depends on each worker's individual preferences for time spent engaged in work and nonwork activities. It is simply not possible to make an a priori prediction, for any person or group of people, as to which effect will dominate. Each worker will make his or her own decision and the cumulative effect of those decisions will determine the ultimate impact of time spent working in the labour market.

It is worth emphasizing that although it is not

possible to predict the impact of wage changes on the basis of theory, empirical studies of physician responses to changes in fee schedules may provide some guidance to policy-makers on what changes to labour supply might be expected.¹⁶ However, because the theoretical prediction is ambiguous, the only way to know for certain what impact a wage change will have on a population of physicians at a given point is to make the change and perform an empirical study to measure the response. The only certainty in such a scenario is that the unit price of work will increase — although workers on the whole may ultimately provide less total work.

There are many challenges for payers resulting from the opposing income and substitution effects associated with wage increases. For example, the Ontario Ministry of Health and Long-Term Care may be concerned about access to a particular group of physicians (e.g., rural family physicians) and want to encourage those physicians to work more hours. The ministry may decide to increase physician fees for this group. If the substitution effect outweighs the income effect, the ministry will have achieved its goal. However, if the reverse is true (the income effect outweighs the substitution effect), then the ministry will have achieved the opposite of its goal: the physicians will work fewer hours and access to the services they provide will be reduced. The 2010 report from the Organisation for Economic Co-operation and Development suggested that income effects could be strong in Canada, noting that

[A]lthough fee-for-service payment should encourage physicians to seek high volumes, income effects arising from high compensation levels as in Canada, and high marginal tax rates, may conversely induce them to limit their effort, lowering productivity.¹

Effects of life stage

It is possible that fee increases might affect the same physician differently at different stages in his or her career. For example, physicians may respond to an increase in their fees over their career by working more (substitution effect). However, physicians may retire early (income effect), or younger physicians may work more hours when faced with a fee increase (substitution effect), while the opposite might be the case for older physicians who prefer to earn the same amount and work fewer hours (income effect). Alternatively, younger physicians may choose to work fewer hours, preferring instead to spend time pursuing other interests, whereas older physicians may have higher personal expenses and thus could choose to work more hours. In-

deed, recent research suggests that physicians are not winding down their practices until well after their 65th birthday — the average age of self-reported retirement is currently 69.2 years.¹⁷

Nonwage benefits

Nonwage benefits (a benefit of financial value that is not connected to base pay) can be a substantial part of total compensation and can also induce income and substitution effects. Importantly, there may be ways to structure the allocation of nonwage benefits to ensure that the substitution effect dominates the income effect (or vice versa), depending on the intended goal.

Two concepts are essential for understanding the effects of nonwage benefits. First, any form of nonwage benefit that reduces the fixed costs of practising medicine (costs a physician incurs that are independent of the number of patients seen or the number of services provided) and is not linked to the number of hours worked creates a pure income effect, with no substitution effect. For example, support for the repayment of loans, support for education, payment of annual fees and contributions to overhead and staffing are all nonwage benefits that reduce the fixed costs of a practice. Second, any form of nonwage benefit that does not reduce the fixed costs of employment and is linked to time worked will induce a pure substitution effect, with no income effect. Examples that create a pure substitution effect include retention bonuses and payments tied to holding an evening or weekend clinic.

Compensation packages

The design of a total compensation package (wages or fees, nonwage benefits) intended to encourage the provision of more services can be complex. As described earlier, nonwage benefits should be linked to time worked or services provided.

Designing an appropriate package can be a challenge when an employer faces an objective beyond increasing the incentive to work, such as attracting physicians to a rural area or encouraging physicians to move to a new model of payment (e.g., capitation). Nonwage benefits that are not related to time worked, such as contributions to lease or loan payments, licensing, staffing or housing, are often used to attract new physicians to a particular jurisdiction. However, these benefits also reduce the fixed costs of a practice and create a pure income effect, thereby inducing fewer hours of work and fewer services provided. Thus, some forms of nonwage compensation produce a trade-off; they encourage new entrants, but they discourage the hours of work or the number of services provided.

Labour economic theory suggests that, when given a choice, workers will be more likely to choose forms of nonwage compensation that are not tied to hours worked (or number of services provided). That is, they will prefer forms of compensation that have a pure income effect, all other things being equal. This association is supported by a series of focus groups conducted in early 2008 and involving Ontario physicians as part of an effort to gather information about what types of nonwage benefits would appeal to them (Appendix 1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.111018/-/DC1). Physicians were presented with typical nonwage benefits commonly considered to attract new entrants or retain current physicians and asked whether the benefit would have the intended effect. The most popular incentive was payment to support education or professional development — including pay for lost income, locum support, reimbursement of tuition and a travel allowance. The least popular incentives were payments tied to return-of-service requirements. Details on the focus groups, including responses and analysis, are presented in Appendix 1.

Conclusion

Understanding and accurately predicting the response of physicians to incentives is essential if governments wish to increase the supply of physician services. Clearly, this issue is complex, and considerations such as political realities, professional development and quality of care play central roles in the design of policy governing the compensation of physicians. Nonetheless, the basic tools of labour economics provide a perspective on how alternative forms of compensation are likely to affect the supply of physician services.

In a review of compensation policies across industries, Baron and Kreps remind us that

[p]eople try to use pay as a scalpel. Because it means different things to different people, it's a broad sword at best.¹⁸

One certainty — borne out in practise — is that compensation policies that maximize the substitution effect, while minimizing the income effect, will achieve the objective of increasing hours of work. Linking compensation to time worked or services provided, as opposed to forms of pay that are unrelated to time worked, will better ensure the goal of increased work hours. Policy-makers should recognize that policies for compensation may result in just what we hope for — or just the opposite. The theory of labour supply provides guidance as we attempt

to address the critical challenge of managing physician supply and work intensity.

References

1. Organisation for Economic Co-operation and Development. Organisation for Economic Co-operation and Development economic surveys — Canada 2010. Paris (France): The Organisation; 2010.
2. Celia M. Exploring a future where primary care is delivered entirely by non-MD health-care providers. *The Medical Post* [Toronto (ON)] 2011 Jan. 10. Available: www.canadianhealthcarenetwork.ca/physicians/inter-professional/what-if/introducing-a-new-series-what-if-13611 (accessed 2011 Jan. 28).
3. Health Council of Canada. How do Canadians rate the health care system? Results from the 2010 Commonwealth Fund International Health Policy Survey. Canadian Health Care Matters, Bulletin 4. Toronto (ON): The Council; 2010.
4. McGuire TG. Physician fees and behavior: implications for structuring a fee schedule. In: Sloan FA, Kasper H, editors. *Incentives and choice in health care*. Boston (MA): MIT Press; 2008. p. 263-88.
5. Deber R, Wodchis W. *Pay for performance in primary care settings: a synthesis for Ontario*. Toronto (ON): Department of Health Policy Management and Evaluation, University of Toronto; 2007.
6. Golden BR, Sloan FA. Physician pay for performance: alternative perspectives. In: Sloan FA, Kasper H, editors. *Incentives and choice in health care*. Boston (MA): MIT Press; 2008. p. 344-77.
7. Duffin J. The impact of single-payer health care on physician income in Canada, 1850–2005. *Am J Public Health* 2011;101:1198-208.
8. Jeon S-H, Hurley J. The relationship between physician hours of work, service volume and service intensity. *Can Public Policy* 2007;(Suppl 1):s17-s29.
9. Jeon S-H, Hurley J. Physician resource planning in Canada: the need for a stronger behavioural foundation. *Can Public Policy* 2010;3:359-75.
10. Crossley T, Hurley J, Jeon S-H. Physician labour supply in Canada: a cohort analysis. *Health Economics* 2009;18:437-56.
11. Jeon S-H, Hurley J. The relationship between physician hours of work, service volume and service intensity. *Can Public Policy* 2007;(Suppl 1):s17-s29.
12. *British Columbia rural physician programs review* [Appendix A]. Vancouver (BC): Harbour Peaks Management; 2008. Available: www.health.gov.bc.ca/pcb/pdf/rural_review_report.pdf (accessed 2011 Sept. 8).
13. Canadian Institute for Health Information. National Physician Database — payments data 2008–2009. Ottawa (ON): The Institute. Available: <https://secure.cihi.ca/estore/productSeries.htm?pc=PCC476> (accessed 2011 Apr. 5).
14. McGuire TG, Pauly MV. Physician response to fee changes with multiple payers. *J Health Econom* 1991;10:385-410.
15. Hurkey J. *Health economics*. Toronto (ON): McGraw Hill-Ryerson; 2010.
16. Jeon S-H, Hurley J. Physician resource planning in Canada: the need for a stronger behavioural foundation. *Can Public Policy* 2010;3:359-75.
17. Canadian Institute for Health Information. Supply, distribution, and migration of Canadian physicians. Ottawa (ON): The Institute; 2009.
18. Baron JN, Kreps DM. *Strategic human resource management*. New York (NY): John Wiley & Sons; 1999. p. 268.

Affiliations: From the University of Toronto and the University Health Network (Golden); the Collaborative for Health Sector Strategy, Rotman School of Management (Golden, Hannam), University of Toronto, Toronto, Ont.; and the Rotman School of Management (Hyatt), University of Toronto, Toronto, Ont.

Contributors: All three authors conceived and designed the article. Doug Hyatt provided the primary expertise in labour economic theory, including interpretation of the results from the focus groups, and Brian Golden and Rosemary Hannam provided expertise in health policy and the application to the health care context based on experience working with the Ontario Physician Services Committee. All three authors drafted and revised the manuscript, and approved the final version submitted for publication. Brian Golden takes responsibility for the paper as a whole.