

Reflections on the general theory of second best at its golden jubilee

Richard G. Lipsey

Published online: 13 June 2007
© Springer Science+Business Media, LLC 2007

Abstract The origin of the second best article is described and criticisms assessed. Distortions making impossible the achievement of either first or second best optima are outlined. Attempts to establish the applicability of first best rules are criticised, as are general rules for making piecemeal efficiency improvements. Both often use models containing empirically invalid assumptions and a selected few of the full set of distortions. Practical policy advice requires more parochial objective functions than community welfare; must rely on formal and appreciative theory, empirical evidence, and large doses of judgment; and should concentrate on making piecemeal improvements in context-specific situations.

Keywords Second best · Piecemeal policies · Context-specific policies · Distortions · Efficiency conditions · Optimality conditions

JEL Classification D60

1 Formal and appreciative theorising

We begin by noting two approaches to theorising and policy advising. *Formal theories* are expressed mathematically, while *appreciative theories* are developed rigorously in verbal terms. The contrast between the two can be seen by considering two types of argument justifying the market economy.

This paper is a revised version of a Key Note Address
Presented to the 62nd Congress of the International Institute of Public Finance.
I am indebted to Robin Boadway, Mark Blaug, Avinash Dixit, and Curtis Eaton for criticisms and suggestions.

R.G. Lipsey (✉)
Simon Fraser University, RR#1 Q70, Bowen Island, BC, V0N 1G0 Canada
e-mail: rlipsey@sfu.ca

The first is the formal proof of the “two fundamental theorems of welfare economics.” The important proposition for what follows is that the perfectly competitive idealisation of the market economy leads in equilibrium to an *efficient* allocation of resources and, given sufficient value judgements (e.g., a social welfare function can be assumed, and theoretical compensation is a valid criterion for judging social gains), to an *optimal* allocation.¹

Although the proof of the efficiency and optimality of the perfectly competitive market is an intellectual triumph, it raises practical problems because the required assumptions do not mirror reality. So, economists must rely on their subjective judgements about the relevance of these theorems to policy. Here opinions vary from ‘highly relevant’ to ‘totally irrelevant’—surely a most unsatisfactory situation.²

The second argument uses appreciative theorising. Two key propositions are (1) the market system co-ordinates economic activity better than any known alternative—not optimally, just better than the alternatives and (2) markets do this *relatively* efficiently by producing prices that are influenced (but not solely determined) by relative scarcities.³

Those who rely on the formal defence typically seek to give policy advice that is derived ‘scientifically’ from theories containing neither value judgements nor guesswork. Those who rely on the informal defence typically deny this possibility, seeking instead to base advice on a combination of formal models, appreciative theorising, empirical knowledge, and a large dose of judgement.⁴ In what follows, examples of each approach are discussed and arguments are advanced for the practicality of the latter alternative.

2 The genesis of the general theory of second best

During my first term as a PhD student at the LSE, I attended a lecture by Dr. Helen Makower on Viner’s *Theory of Customs Unions* (1950). Although the concepts of

¹In public economics, much is made of the distinction between the purely positive concept of a Pareto efficient allocation and the normative one of an optimum allocation, which requires the value judgements referred to in the text. Most of what is mentioned in this paper is applicable to both concepts.

²Blaug (2007) refers to the differences in views over the practical applicability of the welfare theorems as economists’ “intellectual schizophrenia” and goes on to contrast the judgement of Baumol and Akerlof on the real-world inapplicability of the theorems with statements such as Starr’s (1997:151) that they provide “... a significant defence of the market’s resource allocation mechanism” and Mas-Colell et al. (1995: 556) that they offer “... a strong conceptual affirmation of the use of competitive markets, even for dealing with distributional issues.”

³This justification uses two other well-known propositions: compared to its alternatives, a well-functioning market economy with the necessary institutional underpinnings (i) has fewer concentrations of power, involves less coercion, and fewer opportunities for corruption (not ‘none’ just ‘less’ and ‘fewer’) and (ii) is more conducive to growth. Some readers have suggested that the informal justification must have some formal theory behind it. To the contrary, it has been based on a long line of appreciative theorising running from Adam Smith to Milton Freedman and Thomas Schelling and many others who have expounded these four propositions verbally.

⁴For example, Baumol (2002: 143) calls the second theorem a “fairy tale” that should be discarded, or as Mark Blaug (1997: 255) puts it: “... these beautiful theorems are mental exercises without the slightest possibility of ever being practically relevant...”.

trade creation and trade diversion seemed insightful, I thought they neglected the demand side. That night in a flurry of excitement, I developed the proof that the demand side benefits could outweigh the supply side losses of trade diversion, a proof that I subsequently published in Lipsey (1957). Over the next two years, 1953–1955, I worked on customs union theory. My PhD dissertation was accepted in 1958 and subsequently published (Lipsey 1970).⁵

During my first year on the staff, 1955–1956, it occurred to me while reading an article on customs unions by Andrew Ozga (1955) that many economists, myself included, were discovering the same result in different contexts but not recognising its generality. When Kelvin Lancaster entered my room, we discussed my insight. He said it had occurred to him, when reading Samuelson's *Foundations* (1947), that there was no justification for piecemeal policies establishing first best conditions when the fully efficient allocation was not attainable.⁶ Later that day, I expounded my insight to Harry Johnson who encouraged me to develop it. Kelvin invented a proof of the main proposition that afternoon and submitted it to *The Review of Economic Studies*. Harry, who was one of its editors, suggested that Kelvin and I should join forces. He argued that Kelvin's bare proof would probably go unnoticed and to have an impact we must relate our proposition to the existing literature.⁷ He provided several references unknown to me. So the form in which the article finally appeared owed much to Harry.⁸

2.1 What was new?

What I should have known was that James Meade covered many second best issues in his *Trade and Welfare*. But as explained in detail in Lipsey (1997: Introduction), I had neither attended his lectures nor seen the proofs of *Trade and Welfare* that my fellow graduate student, Max Cordon, was reading in 1953–1954. Nonetheless, it may have been in the air where I picked it up unconsciously—in a not unfamiliar story concerning originality.

Neither did I know of the many precursors to my 'original' idea. In the *Palgrave Dictionary* (1987: 280), Peter Bohm states: "This state of affairs in 1956 was somewhat puzzling considering that the Lipsey–Lancaster conclusion was not entirely novel." He notes that both Pareto and Samuelson had observed that if some of the optimum conditions could not be fulfilled, there was no presumption that fulfilling others would improve efficiency or welfare.

⁵I have told the surrounding story in more detail in Lipsey (1997: Introduction: An Intellectual Autobiography).

⁶This is not something I could have done at the time because I had arrived at the LSE to do my PhD totally innocent of anything beyond high-school mathematics and had to teach myself the calculus that I needed for my thesis.

⁷Of the articles subsequently cited in Lipsey and Lancaster (1956), I had read at that time: Viner (1950), Corlett and Hague (1953–1954), Ozga (1955) and Little (1951). I am indebted to Robin Boadway for pointing out the importance in this context of Ramsey's paper on optimal taxes and Hotelling's on monopoly pricing, a precursor of optimal tax theory.

⁸As originally published, the article had an embarrassingly large number of slips in the formal manipulations, particularly in the sections on the nationalised industry and the general proof of the main theorem. I corrected these when the article was reprinted in Lipsey (1997) (although the section numbers were for some unaccountable reason, omitted from that reprint).

All I do know is that in my office that morning in 1954, I got the insight in a blinding flash.

2.2 Some terminology

Factors preventing attaining an efficient resource allocation are variously called ‘constraints’ or ‘distortions.’ Since neither of these terms cover everything that follows, I use the term “sources of divergence,” *sources* for short. I define these as anything that if introduced on its own would prevent the achievement of a perfectly competitive, price-taking equilibrium that was Pareto efficient and otherwise attainable.

Some confusion was caused by our use of terms. A ‘second best situation’ referred to *any* situation in which the first best was unachievable. The ‘second best optimum setting’ for any *source* referred to the setting of that *source* that maximises the value of the objective function, given settings on all the other existing sources. I follow those usages here.

2.3 What we thought it said

Bohm (1987: 281) says L&L “. . . tried to formulate a constraint which would cover most of what the literature had observed as obstacles to achieving a first best Pareto optimum.” We clearly did not think deeply enough about the kinds and origins of our *sources*. But we did believe that their ubiquity made it impossible to fulfil most first-best conditions and that the issues that resulted would be revealed by any *source* that made a first best unattainable.⁹ So our one irremovable *source* stood for the vast number of *sources* that could not be removed in given circumstances. Some criticisms seemed to miss that point.

- Athanasiou (1966) argued that our constraint of a monopolist charging k times marginal cost was unrealistic because the profit maximising relation between marginal cost and price changes as the demand curve shifts. We agree. But the k rule was meant illustratively not realistically.
- The theory lacks generality because it is inapplicable where the economy is separable, for example, no externalities and a subset of traded goods that are neither complements or substitutes¹⁰ (Davis and Whinston 1965). Agreed, if two parts of the economy do not communicate with each other, there can be no second best inter-relations between them. This is what L&L meant to rule out when we assumed a “general equilibrium”—everything depended on everything else. I do not, however agree that this qualification has the effect of “narrowing the domain of the second best conditions. . . [making] the application of second best rules appear manageable. . .” (McKee and West 1981: 439) since I know of few, if any, empirically demonstrated cases of separability.

⁹Indeed most of the then-existing books on welfare economics devoted much space to showing the how impossible it would be to fulfil the efficiency conditions in real world. See, for example, Little (1957) and de Graff (1957).

¹⁰The latter condition requires that the good has a unit elasticity of demand and be additively separable from all other goods in the utility function.

- The theory has been characterised as “...having as its aim, the recommendation of a particular policy to some existing government...” (McKee and West: 441) and has been criticised for positing a benevolent policy maker (Hoff 2001). But our proof says nothing about the motivation of policymakers. It only predicts the outcome of policies adopted in second best situations for any reason.
- “...the policy recommendations depend on the taste/technology specifications of the model and the latter are difficult to assess empirically” (Hoff 2001). We believed that second best conditions would be far too complex to provide realistic policy guides. Hoff’s point reinforces this, as does Athanasiou (1966: 85) when he writes: “...it is hard to see how any of the imperfections current at a time will be so enduring and stiff as to justify an adjustment of the whole system to them.” Indeed, governments do not attempt to achieve either first or second best optima, nor do they often attempt to remove the source of the “distortion;” instead they often attempt to influence its consequences.

2.4 An illustration of what it said

For illustrative purposes in what follows, I summarise the example in Lipsey and Lancaster (1956: Sect. V). Let there be a small country specialised in the production of commodity X and importing commodities Y and Z at fixed prices. Initially assume an equal ad valorem tariff on the imports and no other sources of ‘distortion.’ In this situation, there will be too much consumption of the domestic good X and too little of both imports Y and Z compared with the first best optimum. To consider a customs union, I asked: given that consumption of the domestic good is untaxed while imports of Z are subject to a fixed duty, t_z , what is the second best import duty, t_y , on commodity Y ? What I proved was that for a second best optimum (i) $t_y > t_z$, (ii) $0 < t_y < t_z$, (iii) $t_y < 0$ (a subsidy), according as in consumption Y is (i) complementary with X , (ii) a substitute for both X and Z (iii) complementary with Z .

3 First and second best resource allocations

Next, I argue that the set of realistic policy goals does not include achieving either an economy-wide, Pareto efficient allocation of resources or an economy-wide second best setting for any one *source*.

3.1 Types of *sources*

L&L did not give enough consideration to the nature of *sources*. Some are created by private agents, some are created by the state and some are given exogenously and cannot be altered by policy. It is often argued that exogenously given constraints become part of the first best solution and so can be ignored, leaving only policy-created *sources* to be the subject of second best theory. Although such things as the budget and resource constraints are not *sources*, other exogenous constraints are. For example, because information is costly to acquire, agents remain ‘rationally ignorant’ when the cost of acquisition exceeds its value to the agent. But the state can rationally

acquire the information and provide it to agents at a cost-covering price that is a fraction of each agent's private acquisition cost (assuming that the state's cost is less than the sum of individual agents costs). This example illustrates the important proposition that *sources* that are in the nature of things can lead to inefficient allocations so that there is scope for policy intervention to improve efficiency as long as some of the consequences can be altered. So not all *sources* are policy created, something that is denied by many writers who seek to restore first best conditions as guides for piecemeal policy (e.g., Faith and Thompson 1981; McKee and West 1981 and Hoff 2001).

Next, consider some of the most important real-world *sources*, many of which are not policy created.

3.2 Static considerations¹¹

1. Market structures are rarely competitive enough to make marginal cost equal to price: oligopoly, monopolistic competition and monopoly vastly outnumber cases where firms are price takers. Some price setting behaviour occurs because of technologically determined factors such as scale economies, some because of firm-determined entry barriers and product characteristics¹² and some because of policy.
2. Since most products are differentiated, fixed costs that create significant non-convexities are ubiquitous: entry costs to establish distribution networks, product development costs, and advertising needed to introduce new products.
3. Location in space creates overlapping oligopolies where neither monopolistic nor perfect competition is typically possible (Eaton and Lipsey 1989 and 1997: Introductory Essay). Fixed costs ensure that space is inhabited by "lumpy" firms located at distinct points in space. This implies that free entry will not drive profits to zero (Eaton and Lipsey 1978). Furthermore, the Nash equilibrium under free entry produces a pattern of rectangular markets rather than the efficient pattern of Löschian hexagons (Eaton and Lipsey 1976).
4. Many labour markets are not auction markets. Wages are often payments on implicit long term economic contracts, varying with age. Wages are often signalling devices. Labour markets are often internal, employers promoting existing employees rather than searching outside for better candidates. Even where these, and many other similar forms of behaviour, are efficient responses to non-perfectly competitive circumstances, they upset the Paretian conditions in labour markets.
5. Governments intervene in many markets with such things as rules, regulations quantity restrictions, taxes and subsidies, import tariffs and non-tariff barriers.
6. Incomplete and asymmetric information abounds.

¹¹No doubt, there is some overlap among these items and if I were writing a text book, I would strive to remove them. For present purposes, however, the extent and variety of *sources* is what matters so that any overlap among *sources* does not detract from the general argument.

¹²There is no impersonal market in which the price of a generic version of differentiated products, such as refrigerators, is determined. Individual manufacturers must administer their own prices and take externally determined sales as their market signals. For discussion of the effect of product differentiation on the competitive model see Eaton and Lipsey (1989).

7. Positive and negative externalities are attached to many economic activities.
8. There are many missing markets.
9. One of the foundations of welfare economics, the maximisation of utility functions in which the only arguments are the commodities consumed by the agent in question, is currently being challenged (Layard 2005). Modern research confirms that individuals are social animals and what others do enters into their utility functions in myriad ways. This greatly alters the set of policy changes that can increase welfare.

3.3 Dynamic considerations: endogenous technological change¹³

Innovation typically involves Knightian uncertainty (Knight 1921). Risky events have well-defined probability distributions and expected values. Uncertain events have neither. In a risky situation, two agents desiring the same objective, having the same information, and choosing among the same alternative actions, will make the same choice—the one that maximises the expected value of the outcome. In uncertain situations, however, these two agents may make different choices, neither of which can be shown *ex ante* to be better than the other.

Because firms make R&D choices under uncertainty, there is no unique line of behaviour that maximises their expected profits—if there were, all equally well-informed competing firms seeking the same breakthrough would be doing identical R&D.¹⁴ Thus, firms are better seen as groping into an uncertain future in a profit-seeking manner, rather than maximising the present value of expected future profits. Such groping behaviour makes technological trajectories path dependent and non-unique. An important implication is that the conditions for an efficient allocation of resources cannot even be defined when technology is changing endogenously under conditions of uncertainty—we do not know what allocation will produce the best results, however defined, until *after* the results are in. From this follows that there does not exist a set of scientifically determined optimal public policies with respect any variable such as taxes, subsidies, tariffs, commercial and industrial policy, that influence R&D directly or indirectly.

3.4 Conclusion

We do not have a GE model of an institutionless, fully free market economy with the mix of market forms and that characterises a typical industrialised economy. Thus, there is no compelling theory or evidence to suggest that such economies are statically efficient and many government policies reveal the judgement that the allocation achieved by such markets can be improved. Furthermore, we do not have a model that incorporates the other static *sources* mentioned above. Finally, once we allow for endogenous technological change accomplished under uncertainty, we cannot even define the conditions for an efficient allocation of resources. The upshot is that in

¹³The material in this section is based on Lipsey et al. (2005: Chaps. 2 and 4).

¹⁴Compelling illustrations of the consequences of choice made under uncertainty are seen when competing Japanese and American firms make different R&D decisions although both are searching for the next advance in some product over which they compete. For examples, see Dertouzos et al. (1989).

practical situations, as opposed to theoretical models, we do not know the necessary and sufficient conditions for achieving an economy-wide, first-best allocation of resources.

Achieving an economy-wide second best optimum allocation looks even more difficult than achieving the first best. Without a model of the general equilibrium that contains most let alone all of the above *sources*, we cannot specify the existing situation formally and so cannot calculate the second best optimum setting for any one *source* that is subject to policy change.¹⁵ This is an important point since much of the literature that is critical of second best theory assumes that economists know a distortion when they see one and know that the ideal policy is to remove the distortion directly, something that is necessarily welfare improving only in the imaginary one-distortion world.

This leaves us with two questions: (1) Can we find general conditions for piecemeal improvements in welfare? (2) Can we make piecemeal improvements in specific situations? I consider these questions in Sects. 5 and 6. But first, I must consider some issues that arose from our article but which I could not critique without the material just covered.

4 Subsequent developments

Boadway (1997a: 3) points out, the large literature on optimal taxes, as well as several other issues in public economics, address second best problems. However, I find that when graduate students trained in general theory are asked for policy advice with respect to a taxed market, they typically suggest removing the ‘distortion’ with no consideration of possible second best ramifications.

As discussed by Davis and Whinston (1965 and 1967), two interpretations of second best theory were advanced soon after our article. One, advocated by authors such as McManus (1958–1959) and Bohm (1967), argued that “. . .all relevant economic problems must be solved within the context of a general equilibrium model (Davis and Whinston 1967: 324).” The other, championed by D&W, was that “. . .it is not helpful (and probably not feasible) to view all economic problems within the context of a general equilibrium model (Davis and Whinston 1967: 324).” D&W investigate conditions under which piecemeal policy can apply second best rules to a sub-set of the economy while ignoring the remainder. Lancaster and I saw the message of our article as being close to D&W’s interpretation. However, because of the existence of a wide range of *sources*, most of which are not even considered in existing models, we believed that for all practical purposes global second best maxima were unobtainable and also that the piecemeal application of first best rules, where they could be imposed, had no guarantee of improving efficiency. So what remained for policy was to make piecemeal changes.

There is insufficient space to consider all of the many attempts to show that piecemeal policy based on establishing first best conditions is still relevant. I consider three very different arguments each of which has been widely quoted with approval.

¹⁵Indeed, Boadway reports that “Stiglitz general account of conditions for welfare improvements in multi-consumer situations are technical and not easy to interpret.” And this is in a model that ignores virtually all of the *sources* listed above!

Faith and Thompson (1981) introduce a monopoly as the single ‘distortion’ and argue: “While much of the applied literature appealing to “second best” arguments suggests all sorts of informational limitations on government decisions makers, this is hardly reasonable given that the decision maker must know the $U[t]$ and $T[t]$ functions in order to ascertain the second best policy.”¹⁶ So the only plausible reason for not attacking a distortion directly is that voters do not trust policy makers to adopt the correct policy of a per unit subsidy and a lump sum tax. But a second best set of taxes on the rest of the economy designed to increase the monopolist’s output to the competitive level would also increase its profits, causing a paradox between what was not allowed directly but was allowed through a second best back door. The paradox is removed by adopting a set of taxes and subsidies in the rest of the economy that leaves the monopoly unaffected. This, not surprisingly, calls for zero taxes. So the paradox is removed by applying first best elsewhere, leaving the monopoly unaffected.

My objections are as follows. First, the argument that the only source of persistent ‘distortion’ is fear of perverse behaviour on the part of government requires the unjustified belief that governments understand the general equilibrium of the entire economy. This understanding is required if, as F&T assume, governments can design a global second best policy. If as I argue, policy makers lack the required knowledge, monopolies may be tolerated for reasons other than fear. Second, policy makers are faced with many *sources*, that are not policy induced. When in a many-*source* economy they seek to alter the behaviour of a monopoly, they do so in a second best setting where making marginal cost equal price (whether by acting directly on the firm or indirectly by altering things that impinge on its decision) is not necessarily the optimal policy.

The second example alleges that “The Lipsey–Lancaster formulation . . . begs the question of why the distortion that makes the economy second best is out of bounds for policy makers. . . .” (Hoff 2001). “Bohm (1967) recognised that if the distortion were escapable, that is accessible to direct policy action, the second best policy would, by definition, vanish” (McKee and West 1981: 442).¹⁷ In a world of many *sources* of various kinds, attacking all *sources* directly is not a feasible policy and removing some does raise second best issues. The same comment applies to Bhagwati (1971: 77) who argues that optimal policy is to attack the source of the distortion directly.

Another argument for a presumption in favour of first best policies has been advanced by Ng (1977) and quoted favourably by several authors (e.g., Hoff 2001; Bohm 1987).¹⁸ He assumes a parametric ‘constraint,’ c_p , and one that can be varied by policy, c_v , plus a welfare function that is concave plotted against the degree of

¹⁶This argument puts Faith and Thompson in the same camp as McManus and Bohm in holding that second best policies consist of solving the GE of the entire economy to produce the second best optimum.

¹⁷Space limitations prevent me from considering in detail McKee and West’s (1981) critique of second best. They are in the camp who believes that all relevant distortions must be policy induced and they develop a critique based on the idea that many of these are optimal outcomes of the political market place. Possibly, but many sources are not policy induced and of those at least some are welfare reducing.

¹⁸Ng calls a situation in which the second best optimum cannot be achieved and small improvements in welfare are sought “third best.” We call this piecemeal second best policy (as opposed to optimum second best policies) but as long as we know what is being done, the words used to describe it are immaterial.

departure from the first best rule by either constraint because "... it is reasonable to expect that as we diverge more and more from the first-best rule, the marginal damage increases" (Ng 1977: 2–3). First, establish a first best by setting both constraints at zero. Then, set the parametric constraint at a non-zero level. This shifts the welfare function $W = W(\bar{c}_p, c_v)$ but the policy makers do not know which way it has shifted. If altering c_v moves the economy towards its second best optimum, the absolute value of the welfare gain is less, by the convexity assumption, than the absolute value of the welfare loss if the alteration moves the economy away from its second best optimum. So the expected value of the utility function is maximised by holding to the first best rule, $c_v = 0$.

Ng takes as "reasonable" the convexity of the welfare function plotted against the degree of departure from the first best solution. But this is by no means necessary. As an illustration, consider the example outlined in Sect. 2.4 above with the addition of a Cobb Douglas utility function as analysed in Lipsey (1970), $U = X^\alpha Y^\beta Z^\gamma$. The implied welfare function $W = W(t_y, t_z)$ is maximised for $t_y = t_z = 0$.¹⁹ However, the function $W = W(t_y, \bar{t}_z)$ is not everywhere convex. While its first derivative is everywhere negative, its second derivative shifts from negative to positive as t_y increases beyond a point of inflection. This should not be surprising since it is impossible to drive welfare to zero with any finite level of tax given these tastes. So while as t_y is increased over some range starting from zero the W function does get steeper, eventually its steepness begins to decrease. Hence for this case Ng's argument is correct for some small enough values of \bar{t}_z but is reversed for larger values since the expected value of a small increase in t_y from zero is then positive.

5 Are there general policy rules for piecemeal improvements?

My answer to the question posed in the heading is "no". As Boadway (1997b: 756–657) puts it: "...though duality theory operationalised second-best theory, the results were not encouraging: simple piecemeal policy rules do not hold in a distorted, second-best world except under extremely unlikely combinations of taste and technology." And this is said without taking into account most of the *sources* that I outlined in Sect. 3.

Nonetheless, some economists have tried to develop useful *general rules* for making piecemeal improvements. To consider these, I first look at Harberger's classic 1971 article, where he argues that second best policies are easier to establish than L&L maintained. In a space for the set of all (economic) activities he plots the subset of those that are affected by significant 'distortions' and the subset of activities that would be affected if public policy alters one of these. He argues that the intersection of these sets is all that matters for small changes in public policy because a marginal policy-induced change in any non-distorted activity has no significant effect on welfare. He states that the area of this intersection is small enough "...to dispel

¹⁹For the case of $\alpha = \beta = \gamma = 1/3$, the welfare function is $U = [\frac{A^3 i^2 j^2}{p_y p_z (i+j+i+j)^3}]^{1/3}$ (Lipsey 1970: page 46, (1)) where the $i = (t_x + 1)$, $j = (t_z + 1)$, A is the country's endowment of X and p_y and p_z are the prices of the two imports the domestic good, p_x being the numeraire.

any thoughts that the job of incorporating general-equilibrium aspects is so big as to be effectively hopeless.” (791) My response is twofold. First, the material in Sect. 3 makes me sceptical of Harberger’s judgement that the proportion of the total space affected by ‘distortions’ (my *sources*) is small. As a first approximation, I would say it was close to 100%. Second, I know of no evidence that the space of activities that is significantly affected by the typical new policy is small. After all, sectors everywhere can be affected by a ‘distortion’ anywhere and to say that only a few markets will be affected by changing a *source* in one is an empirical judgement whose substantiation requires evidence gathered case by case.

I now consider some specific generalisations. Most of these, as well as the three more general arguments that I criticised in the previous section, are subject to three types of objections.

Type 1 objections: Only one type of *source* is considered, such as taxes, and then usually only two items from this *source*, one that is given and one that can be varied by policy. No one knows if the results will stand in models with more items from the one type of *source* plus items from other *sources*.

Type 2 objections: Many of the propositions are based on restrictive assumptions not found in reality and so provide no obvious guide for practical policy.

Type 3 objections: The possible effect on technological change is ignored—a serious shortcoming since small induced changes in the growth rate can have large cumulative effects on GDP.

Many of the specific generalisations that follow are corollaries of the one proposition stated in Sect. 2.4 above: when all goods are substitutes for each other and are subject to a set of indirect taxes, the second best value of any one tax, given fixed values of all the others, is lower than the highest tax and higher than the lowest tax. We use this example in what follows

1. *Lowering the highest distortion must be welfare increasing.* In the three commodity case if $t_y > t_z$ reducing t_y will necessarily increase welfare only if Y and X are substitutes in consumption. If they are complementary, reducing the value of t_y in the neighbourhood of the value of t_z will necessarily reduce welfare.
2. *Two small distortions are better than one large one.* This is not necessarily true on its own terms (ignoring all three types of objections, which also apply). We saw when discussing Ng, that if the utility function is Cobb–Douglas, the function showing how welfare changes with the variable constraint goes through a point of inflection. It follows that for some high enough tax rates, situation 1 in which ($t_{y1} = t_{z1} > 0$) will yield less welfare than situation 2 in which ($t_{y2} = 2t_{y1}$, $t_z = 0$).²⁰
3. *The optimal level of a policy reduced distortion will fall short of correcting the distortion fully.*²¹ Again this is not true if Y and Z are complementary in consumption, in which case not only is each reduction of the tariff on Y down to zero welfare improving, the second best optimum requires $t_y < 0$.

²⁰For example for the utility function given in the previous footnote the utility associated with a uniform 50% tariff on Y and Z is less than the utility associated with a 100% tariff on Y and a zero tariff on Z .

²¹Hoff (2001) attributes this to Bhagwati (1971). Hoff does not mention the qualification that the goods must be substitutes while Bhagwati does.

4. *A small dose of a policy that has some effect on the distorted margin is better than no policy at all, because the initial marginal gain from mitigating the distorted market is of first order while the initial welfare cost from introducing the new distortion is of second order.* Bhagwati et al. (1969: 1009) This applies only to the case where the policy introduces a distortion that was previously zero. Given the *sources* discussed in the previous section, it is hard to identify a real world market that does not already contain some *sources* such as marginal cost does not equal price.
5. *Starting from an arbitrary set of tax rates and reducing the highest and raising the lowest will be welfare increasing if tax receipts rise in both markets and if both goods are substitutes for the aggregate of all goods.* (Hatta 1986: 105) This theorem not only requires the conditions stated in point 4 but also that there are no other types of taxes (and no other *sources* since their possible effects on the theorem are not investigated).
6. *Starting from a tax-distorted situation, a proportional reduction in the size of all distortions will generally raise welfare.* (Atkinson and Stiglitz 1976 as quoted by Boadway 1997a: 11)

All of these, and virtually all other similar propositions in public economics, are open to type 1 objections in that all types of *sources* other than the one being considered are ignored. For example in point 6, how do we measure an “equal reduction” in each of the many items in each of the types of *sources*?

Almost all of these generalisations are open to type 2 objections in that all but number 6 require restrictive assumptions about substitutability. Other generalisation not listed here are also open to type 2 objections in that they make empirically unsupported assumptions about separability in the utility function.

All the generalisations are also subject to type 3 objections. For example, with respect number 2, two monopolistically competitive firms, each with only a small divergence between marginal cost and price may not be better than one oligopoly with a large divergence yielding high profits used to finance R&D?

6 Context specific piecemeal policies

Can we use welfare economics to derive useful policies that are not open to the objections already stated? It is more obviously useful where the objectives are more parochial than maximising the welfare of the whole society, the objective function is clearly specified, spillovers may not matter (or can be roughly taken into account), partial equilibrium analysis is useful and both direct and the most obvious indirect effects can be measured.

Harberger seems to agree when he writes (1971: 795) that rather than telling how to reach global second or first best optima “[t]he practitioner. . . is more likely to be asked which of two alternative agricultural programs is better, or what resource-allocation costs a given tax increases involves, or whether a certain bridge is worth its cost.” Here are further examples.

Let the task be to reduce noxious emissions by a given amount at the least cost. There may be indirect effects, but either these do not matter to policy makers or

they can be roughly estimated. Unlike theories that predict how to maximise society's welfare, the predictions of these theories concerning the relative advantages of alternatives can be tested against direct measurements of the results.

For years Ontario Hydro insisted on pricing electricity at the average cost of production by low-cost hydro electric, high-cost coal fired and medium-cost atomic energy. New plants all produced at costs higher than this average but were built whenever there was excess demand at the going price although customers were unwilling to pay the marginal cost of producing electricity with them.

Many principal agent problems are fairly contained. The principal is interested in getting good work from his agents. Theories about alternative methods of doing this can be tested against observable results. There may be second best side effects but that is not the issue. The economist's brief is to get what the principal wants at the least cost.

In these and most similar cases, policy advisers need a good knowledge of 'Harberger rectangles,' the theory related to such things as pollution permits, and a general understanding of second best issues as a caution, while highly sophisticated formal GE theory adds little value to the tool kit. When such non-GE methods can be used and how many indirect effects to consider are a judgements based on appreciative theorising—not something that can be established conclusively through formal modelling.

But although there can be few if any valid general *policies*, much of the useful work emanating from second best theory has been to develop scientific *approaches* that can be generally useful when adapted to local circumstances. For example, cost-benefit analysis is an area where scientific principles based on second-best analysis have been widely and successfully applied. The notion of shadow pricing of project inputs and outputs to take account of market 'distortions' is a general principle. Nonetheless, because these techniques do not take account of all *sources*, a large element of judgement is needed whenever they are used to derive specific policy advice.

7 Conclusion

In all practical circumstances, economists investigate policy issues using methods that omit a potentially significant subset of *sources*. Thus, we must of necessity make personal judgements about the applicability of such models when predicting where piecemeal, second-best improvements are possible. This is one of the many reasons why policy advice must use a mixture of formal modelling, appreciative theorising, relevant evidence and an inevitable amount of judgement—and why it must be context specific (i.e., there are few practical generalisations that apply to each and every set of items in each and every *source*). The task is easier if the objective function is more circumscribed than the whole society's welfare. Although this may be obvious to economists with policy experience, it is not a warning typically emphasised in public economics texts. These mainly give students rigorous proofs of propositions that appear to be policy relevant and universally applicable, but are open to the types of objections stated earlier. Much of the literature that I have surveyed (and there is

much more than I have had space to review) directly objects to these second best conclusions arguing that scientifically derived first best conditions are applicable to real policy situations.

Finally, I offer two replies to the allegation that second best theory provides justification for just about any crazy interventionist policy.

First best theory is also often used perversely as when it is used to justify the one-size-fits-all policy advice that often does more harm than good. For example, it is counter productive to advise a developing country to remove its “distorting” subsidies on sugar production when the result is the elimination of its sugar industry by subsidised US production while its former workers remain unemployed.²² The best response to these one-size-fits-all policy advisors is to observe that they are operating in a second best situation and that, since the specific *sources* vary temporally and spatially, the policies needed for piecemeal improvements also vary.

My second reply is that the ‘anything goes’ implication seems persuasive because it is viewed from the aspect of the formal defence of the price system and the scientific approach to policy advice. When the intuitive appreciative approach is used, both to defend the market economy and to assess policies, it is easier to spot counter productive advice. According to the appreciative approach, since the competitive market economy is the best known method of allocating resources, departures from it through either public policy or private behaviour, are regarded as *prima facie* undesirable, unless justified by well-reasoned and persuasive arguments. This is sufficient to rule out the kinds of massive tariffs that used to be found in many developing countries, but not to rule out carefully designed, administered and sunsetted infant industry tariffs. It also rules out private sector actions in restraint of competition, again unless there are very good reasons for them. Highly elaborate theory is not necessary in these cases and many others like them. What is needed is a good appreciative understanding of how the price system works, as well as understanding the cautionary warning from second best theory that any policy may have unexpected and undesirable consequences in apparently unrelated parts of the economy that need to be watched for and mitigated where necessary. Useful piecemeal policy advising is not impossible; neither can it be determined purely scientifically; instead it is an art, assisted by good economics, both theoretical and empirical.

References

- Athanasiou, L. (1966). Some notes on the theory of second best. *Oxford Economic Papers*, 18(1), 83–87, March.
- Atkinson, A. B., & Stiglitz, J. E. (1976). The design of the tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6, 55–75.
- Baumol, W. J. (2002). *The free market innovation machine*. Princeton: Princeton University Press.
- Bhagwati, J. N. (1971). The generalized theory of distortions and welfare. In J. N. Bhagwati, R. W. Jones, R. Mundell, & J. Vanek (Eds.), *Trade, balance of payments and growth: papers in international economics in honor of V. Charles P. Kindleberger* (pp. 69–90). Amsterdam: North-Holland.
- Bhagwati, J. N., Ramaswami, V. K., & Srinivasan, T. N. (1969). Domestic distortions, tariffs, and the theory of optimum subsidy: some further results. *Journal of Political Economy*, 77, 1005–1019.

²²For examples of this sort of misguided policy see Griffiths (2003).

- Blaug, M. (1997). Competition as an end-state and competition as a process. In *Trade technology and economics: essays in honour of Richard G. Lipsey* (pp. 241–261). Cheltenham: Edward Elgar.
- Blaug, M. (2007). The fundamental theorems of modern welfare economics, historically contemplated. *History of Political Economy*, 32(2), 186–207.
- Boadway, R. (1997a). The role of second-best theory in public economics. In B. C. Eaton, & R. G. Harris (Eds.), *Trade, technology and economics essays in honour of Richard G. Lipsey* (pp. 3–25). Cheltenham: Edward Elgar.
- Boadway, R. (1997b). Public economics and the theory of public policy. *The Canadian Journal of Economics*, 30, 753–772.
- Bohm, P. (1967). On the theory of ‘second best’. *Review of Economic Studies*, 34, 301–314.
- Bohm, P. (1987). Second best. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new palgrave, a dictionary of economics* (pp. 280–283). London: Macmillan.
- Corlett, W. J., & Hague, D. C. (1953–1954). Complementarity and the excess burden of taxation. *The Review of Economic Studies*, 21, 21–30.
- Davis, O. A., & Whinston, A. B. (1965). Welfare economics and the theory of second best. *Review of Economic Studies*, 32, 1–14.
- Davis, O. A., & Whinston, A. B. (1967). Piecemeal policy in the theory of second best. *Review of Economic Studies*, 34, 323–331.
- de Graff, J. (1957). *Theoretical welfare economics*. Cambridge: Cambridge University Press.
- Dertouzos, M. L., Lester, R., & Solow, R. (1989). *Made in America: regaining the productive edge*. London: MIT Press.
- Eaton, B. C., & Lipsey, R. G. (1976). The non-uniqueness of equilibrium in the Loschian model of spatial location. *American Economic Review*, 66, 77–93.
- Eaton, B. C., & Lipsey, R. G. (1978). Freedom of entry and the existence of pure profit. *Economic Journal*, 88, 455–469.
- Eaton, B. C., & Lipsey, R. G. (1989). Product differentiation. In R. Schmalensee, & R. Willig (Eds.), *Handbook of industrial organization* (pp. 725–768). Amsterdam: North Holland.
- Eaton, B. C., & Lipsey, R. G. (1997). *On the foundations of monopolistic competition and economic geography: the selected essays of B. Curtis Eaton and Richard G. Lipsey*. Cheltenham: Edward Elgar.
- Faith, R. L., & Thompson, E. A. (1981). A paradox in the theory of second best. *Economic Inquiry*, XIX, 235–244.
- Griffiths, P. (2003). *The economist's tale: a consultant encounters hunger and the World Bank*. New York: ZED Books.
- Harberger, A. C. (1971). Three basic postulates for applied welfare economics: an interpretive essay. *Journal of Economic Literature*, 9, 785–797.
- Hatta, T. (1986). Welfare effects of changing commodity tax rates towards uniformity. *Journal of Public Economics*, 29, 99–112.
- Hoff, K. (2001). Second and third best theories. In J. Michie (Ed.), *Reader's guide to the social sciences*. London: Fitzroy Dearborn.
- Knight, F. H. (1921). *Risk, uncertainty and profit*. New York: Houghton Mifflin Co.
- Layard, R. (2005). *Happiness: lessons for a new science*. London: Penguin Books.
- Lipsey, R. G. (1957). The theory of customs unions: trade diversion and welfare. *Economica*, 24, 40–46, February.
- Lipsey, R. G. (1970). *The theory of customs unions: a general equilibrium analysis*. London: Weidenfeld & Nicolson. A late publication of the 1957 Ph.D. thesis by the same name.
- Lipsey, R. G. (1997). *The selected essays of Richard Lipsey: Volume I: microeconomics, growth and political economy*. Cheltenham: Edward Elgar.
- Lipsey, R. G., & Lancaster, K. (1956). The general theory of second best. *The Review of Economic Studies*, 24, 11–32. Reprinted with errors corrected in Lipsey 1997.
- Lipsey, R. G., Carlaw, K. I., & Bekar, C. (2005). *Economic transformations: general purpose technologies and long-term economic growth*. Oxford: Oxford University Press.
- Little, I. M. D. (1951). Direct versus indirect taxes. *The Economic Journal*, 61(243), 577–584, September.
- Little, I. M. D. (1957). *A critique of welfare economics* (2nd ed.). Oxford: Clarendon Press.
- Mas-Colell, A., Whinston, M.D., & Green, R.J. (1995). *Microeconomic theory*. New York: Oxford University Press.
- McKee, M., & West, E. G. (1981). The theory of second best: a solution in search of a problem. *Economic Inquiry*, XIX, 436–448.
- McManus, M. (1958–1959). Comments on the general theory of second best. *Review of Economic Studies*, 26, 209–224.

- Ng, Y.-K. (1977). Towards a theory of third-best. *Public Finance*, 32(1), 1–13.
- Ozga, S. A. (1955). An essay in the theory of tariffs. *Journal of Political Economy*, 63, 489–499, December.
- Samuelson, P. A. (1947). *Foundations of economic analysis*. Cambridge: Harvard University Press.
- Starr, R. M. (1997). *General equilibrium theory: an introduction*. Cambridge: Cambridge University Press.
- Viner, J. (1950). *The customs union issue*. New York: Carnegie Endowment of International Peace.