Welfare Analysis: 
Bridging the Partial and General Equilibrium Divide for Policy Analysis

Scott Farrow and Adam Rose

1. Introduction

There are two schools of practice for applied welfare analysis, partial and general equilibrium analysis. While they have some theory and literature in common, each school has its own additional literature and practitioners, with little communication between the two. Some of the different perspectives between the schools and their historical practice is codified by the US Office of Management and Budget (OMB) for both regulatory analysis and benefit-cost analysis in general. While this note seeks to bridge the practitioner’s divide between partial and general equilibrium, an additional purpose is to review the basis for excluding secondary or indirect effects in OMB guidance for the benefit-cost analysis of government investments, policies and rules. The article is based on a literature review where specific citations are provided on the more technical details from sources such as Mas-Colell, Whinston and Greene (1995) Acemoglu (2009) and Varian (1992).

Partial equilibrium (PE) welfare analysis implemented through benefit-cost analysis is the school taught regularly in undergraduate and Master’s economics and policy courses in benefit-cost analysis using texts such as Boardman et al. (2011), Zerbe and Dively (1994), and Bellinger (2007). Such books are almost if not entirely focused on the PE framing and methodology, evaluating direct effects in a limited number of settings or providing for limited cases where the PE approach is applied to several markets. That approach is also institutionalized in US Government practice through Executive Orders for the analysis of major regulations and other

---

1 The authors are, respectively, professor and faculty affiliate of the National Center for the Risk and Economic Analysis of Terrorism Events (CREATE); and research professor, Sol Price School of Public Policy and faculty affiliate of CREATE. Appreciation is extended to the Department of Homeland Security for funding support. This version June 30, 2016.
applications using benefit-cost analysis (OMB, 1982; 2003). The Executive Orders preclude consideration of “secondary” or “multiplier” effects (OMB 1982)\(^2\), although a limited category exists for specific PE elements of “ancillary” benefits or costs unrelated to but caused by the “direct” effects of a regulation (OMB, 2003)\(^3\).

In contrast, a General Equilibrium (GE) approach is more widely accepted in Europe (Florio, 2014; Dreze and Stern, 1987) and in areas of application that tend to cross market boundaries such as macroeconomic growth, international trade, taxation, and major terrorism events (Dixon and Jorgenson, 2013). The GE framing involves a high level of aggregation but models direct and indirect effects transmitted through a chosen number of input and output markets along with the expenditures of governments.

Several decades ago there may have been more reason for such methodological segregation when the conceptual foundations for partial and general equilibrium analysis were more distinct. However, the wide-spread use of neo-classical micro foundations for general equilibrium modeling suggests that this may be a time for re-appraisal of practice. This survey explores the metrics used for PE and GE welfare analyses and their key assumption in Section 2. Numerous variations exist in empirical practice, so what is reviewed here is subjectively focused on “standard” (versus “frontier”) practice (Farrow and Zerbe, 2013). Readers are assumed familiar with the basic distinctions between equivalent and compensating variations (EV and CV respectively) but key issues will be summarized. The primary focus is on static PE and GE models acknowledging the additional extensions in both metrics and estimation procedures for dynamic, stochastic and behavioral models (e.g. Acemoglu, 2009; Bernheim and Rangel, 2009). Section 3 considers the current relevance of the OMB proscription against GE approaches and concludes.

\(^2\) OMB (1992) defines multiplier as “the ratio between the direct effect on output or employment and the full effect, including the effects of second order rounds or spending. Multiplier effects greater than 1.0 require the existence of involuntary unemployment.” Secondary is defined in context as “Employment or output multipliers that purport to measure the secondary effects of government expenditures on employment and output should not be included in measured social benefits or costs.” (OMB, 1992)

\(^3\) “An ancillary benefit is a favorable impact of the rule that is typically unrelated or secondary to the statutory purpose of the rulemaking (e.g., reduced refinery emissions due to more stringent fuel economy standards for light trucks) while a countervailing risk is an adverse economic, health, safety, or environmental consequence that occurs due to a rule and is not already accounted for in the direct cost of the rule.” OMB (2001)
To foreshadow the conclusion, modern PE and GE models share welfare metrics but differ on the maintained hypotheses of price and other linkages among markets. Both models can provide “shadow prices” for variables of concern, with GE models better suited to the complex social policy issues that challenge PE modeling. At the same time, GE models transmit the smallest of shocks, dampened or amplified, throughout the system including effects on input markets such as labor supply. Which model is more liable to ex-post error is little studied. At the same time, the Government’s negation of “multiplier” effects seems to be an artifact of historically simpler models.

2. Welfare Metrics and Assumptions

Applied welfare analysis or as frequently called--benefit-cost analysis—seeks to answer the question when is society’s welfare improved, or even maximized, by taking an action such as investments in education, highways, or alternative policies.

While not underestimating the ability of economist’s to disagree, if welfare analysis focused solely on an individual consumer or producer, then the distinguishing elements among compensating and equivalent variation and consumer surplus can be clearly delineated (Mas-Colell, Whinston and Green, 1995, p. 80-85). Fundamental is their equivalency when there are no income (equivalently, wealth) effects (Mas-Colell, Whinston and Green, 1995, p. 24, 83).

In earlier eras up through the 1980s, distinct methodologies separated micro and macro economics. Micro built a sequence of models beginning with individual actors such as consumers and producers, built to a market level, then a multi-market level. Models tended to use comparative static analysis evaluating changes in discrete equilibria. In that earlier era macroeconomics focused on aggregated components such as the consumption and investment functions. In the 1980s, macroeconomic models became more explicitly built on the same micro-economic foundations aggregating up through individual actors, to markets, and to economy-wide analyses and typically with more attention to dynamic processes (Acemoglu, 2009, chapter 5).
Figure 1 below illustrates the typical tiered sequence of aggregation to report a welfare metric in applied analysis. Aggregation, in the economic sense of being composed of earlier elements, is indicated by a blue diamond. With PE, the aggregation typically stops with one or a few markets. With GE, the level of market detail may be coarser (hence not necessarily an individual market) but more frequently some degree of aggregation into product classes, ultimately equilibrated and aggregated in an economy wide analysis. The following sections will elaborate on these sequences. In contrast with the dominant practice in the United States, the GE sequence will be the focus and PE will then be defined as restrictions, or caveats, about the assumptions used for the GE analysis.

### Aggregating Welfare Metrics

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Individual actors</th>
<th>Individual Market</th>
<th>Multi-Market</th>
<th>Economy-wide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial Equilibrium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Equilibrium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Indicates aggregation such that the latter term is composed of elements of the former.

#### 2.1. General and Partial Equilibrium Welfare Metrics and Assumptions

The welfare metrics in contemporary GE studies are designed to measure the monetary value of a change in position, such as a change in utility for a consumer, which is then aggregated across consumers and at least two markets. The consumer is assumed to follow the rationality assumptions of neoclassical economics. The three metrics in common use are: 1) equivalent variation (EV), 2) compensating variation (CV), and 3) surplus (S). These are usually developed in detail for consumers but can be applied to producers and factor suppliers (Just, Hueth, Schmitz, 2004).

Substantial intellectual effort has gone into distinguishing EV, CV and S. The distinguishing characteristic of EV and CV lies in the reference point for comparison, whether the initial
condition (EV) or the resulting condition (CV). The two measures are equivalent for consumers and equal to surplus (S) when there is no income (wealth) effect of a price change (Mas-Colell, Whinston and Green, 1995, p. 83; Varian, p. 163) but otherwise differ. Applied studies typically, if implicitly, assume no wealth effect (a constancy of the marginal utility of income for an individual consumer and across consumers) and hence the equivalency of EV, CV and S measures. To the extent aggregation occurs—with more on the subject below—not only individual but aggregate measures of EV, CV and S can be estimated and are equivalent when there is no wealth effect. GE modelers have also gone into substantial depth on the decomposition of welfare effects, particularly in regard to tax effects and international trade. Under various assumptions, one can decompose the total welfare effect into economically meaningful components such as a tax interaction effect (Shoven and Whaley, 1992), a “commodity terms-of-trade” effect (Burfisher, 2011), an “endowment” effect, and so on (Huff and Hertel, 2001; Hanslow, 2000).

Additional metrics are used in GE analysis that have a welfare interpretation only under increasingly strong assumptions. One additional GE welfare metric is a revealed preference aggregate approximations to EV and CV variously called Laspeyres and Paasche cost difference or over and under measures (Dixon and Rimmer, 2002; Ng, 1980). These measures, given microeconomic assumptions of exhaustion of budget and macroeconomic closure rules that government and savings are returned to households, are approximations of real national consumption at initial or post-change prices. However, such market aggregates may omit non-market activities or externalities which are often important to the policy issue at hand.

A change in these welfare measures due to an economic shock is a shadow price, the welfare value of a change due to the shock of interest. At the macroeconomic level, such shadow prices appear to be seldom reported but reflect the macroeconomic change in welfare per unit of shock, whether it be a price or a quantity shock. The microeconomic literature, in its search for “plug-in values” has developed various unit welfare measures or welfare measures as a function of conditioning factors such as socioeconomic or other factors (Boardman, et al., 2011).
Initially the choice of reference point for the welfare measure seemed arbitrary but when there are multiple alternatives, then using the initial basepoint through EV seems appropriate (Varian, 1992). However, work by behavioral economists highlights the importance of the reference point in regards to gains and losses (e.g., Knetsch, Riyanto and Zong, 2012; Brennan, 2016) and in regard to other departures from “rationality” such as the choices of addicts (Bernheim and Rangel, 2009; Boardman, et al., 2011). Figure 1 demonstrates the way in which the reference point and the associated EV or CV changes whether the reference point implies a gain or a loss independent of whether risk is involved. At the same time, Brennan (2016) points out the observational equivalency of this asymmetry with modified utility functions in a neo-classical framework.

Source: Knetsch, Riyanto and Zong (2012)

2.1.1. Aggregation

Aggregation adds layers of assumptions to the analysis of a single actor and the schools start to diverge. Individual actors and differing commodities are typically aggregated. GE models typically aggregate activity into a few or even hundreds of markets, which can be national or regional in scope. Must information on all the heterogeneous actors and each individual market
be carried through such an analysis, or under what conditions can there be a more parsimonious but micro-economically accurate representation of the aggregate?

The existence of income (wealth) effects is a confounding factor in aggregation to a single market. Consider if demands are heterogeneously shifted by changes in income and a policy changes the income distribution. Then information on the heterogeneous nature of consumers (and other actors) would be necessary (Acemoglu, p. 150; Mas-Colell, Whinston and Green, 1995, p. 106). Alternatively, if demand functions (derived from appropriate utility or indirect utility functions) are linear in income with a common coefficient on income across actors, then the members of that market can be represented by an aggregate, representative economic actor (Mas-Colell, Whinston and Green, 1995, p. 107; Acemoglu, 2009, p. 151; Varian, 1992, p. 169). Such functions are said to have a Gorman Polar form (Gorman, 1961). Not all functional forms for utility and implied demand are of this form. The popularity of CES and Linear Expenditure Systems may be in part due to their meeting these conditions (Hallam, undated).

But where does concern for a Social Welfare Function enter? Although standard practice uses the equally weighted sum of individual utility as a social welfare function, long-standing concern for differential weighting of individuals exist allowing unequal weights or differing functional forms (Loomis, 2011). Consequent to the (positive) concern for an aggregate demand is the concern whether aggregation over consumers has a welfare implication consistent with some welfare function. The Gorman form, with its fixed coefficient on income or wealth, implies strong normative properties such that aggregation is relevant for welfare evaluation with any form of wealth distribution (Mas-Colell, Whinston and Green, 1995, p. 119). Further, if wealth is distributed optimally prior to any allocation, perhaps as a result of political rules, then aggregation based on Gorman forms for indirect utility imply aggregate welfare measures for any social welfare function (Mas-Colell, Whinston and Green, 1995, p. 119). Other conditions may of course occur, such that an aggregate demand does not exist; or that aggregate demand exists but it does not have welfare implications (Mas-Colell, Whinston and Green, 1995).

The assumption of a representative consumer for welfare analysis is more often explicit in GE modeling and implicit in PE modeling when market level data are used. PE models occasionally
use explicit aggregation of micro-outcomes in place of a representative consumer and frontier analyses may use more complex aggregation than standard practice. However, at least as far back as Samuelson (1947) and Samuelson and Swamy (1974) there is concern with the positive (objective) consistency of assuming a constant marginal utility of income for aggregation. None-the-less, standard practice for both GE and PE is to aggregate micro-data while making this assumption.

In contrast, as there is no income effect on the production side comparable to that on the supply side, supply functions may be aggregated in the absence of externalities and imperfect competition (Acemoglu, 2009; p. 158). Of course the assumption of perfect competition and no externalities is a significant abstraction for applied work where such concerns motivate much of the policy interest in welfare analysis.

A second type of aggregation is across multiple markets such as “food” or “all other commodities”. This type of aggregation can be justified through separability restrictions or by the commonality of price movements whether deterministic or subject to a random error (Lewbell, 1996). Thus GE models, whatever their number of final markets, involves some commodity aggregation as do PE models. As succinctly summarized by Miller (undated, p. 98-101), such aggregation is also central to PE approaches, where an implicit assumption may model the market of interest and “all other” markets.

A related issue of aggregation across markets is the extent of economic activity aggregated. It is now well understood that there are many non-market activities of importance to the economy and markets in which distortions such as externalities, taxes or market power exist. Many benefit-cost studies focus on policies to address such market distortions such as unpriced or incorrectly priced pollution or non-market behavior such as criminal and recreational choices. Any model is an abstraction, but the change in welfare estimated from any of the measures can be imprecise or incomplete if relevant non-market activities or market distortions are omitted. Errors can result both in direct estimation as in a PE analysis, as well as in a GE analysis.
The third type of aggregation may be thought of as the way in which changes in one market are determined or combined with other markets to define an economy wide solution. It almost goes without saying that a competitive market (without taxation) defines an equilibrium where the supply price is equal to the demand price and the quantity supplied is equal to the quantity demanded. These equilibrium or “closure” assumptions reduce a four variable system of equations down to two equations in two unknowns, which are solvable depending on the functional form. Such assumptions are widely used in both PE and GE. More recently, a significant area of application of CGE models has been to place limitations on the availability of critical inputs to the production process, such as electricity and water services, caused by a natural disaster or terrorist attack. Constraints are placed on these inputs, so that the market equilibrium deviates from the unrestricted counterpart (e.g., Rose and Liao, 2005; Rose et al., 2009, Wing et al., 2015).4

The complexity of GE analysis typically requires additional closure rules regarding major account balances in a macro economy to solve the system of equations. The main consideration is whether one assumes these accounts are in equilibrium or disequilibrium, though this often couched in terms of which variables are exogenous and which are endogenous as noted below (Burfisher, 2011). Major accounts or markets to which this applies include: labor market, markets for traded commodities, and investment and savings, often referred to as “macroclosure”.

The most oft-considered closure rule relates to the labor market. One major approach is often termed the “Keynesian-close rule”, which allows for an underemployment equilibrium by the device of fixing (holding constant) the wage rate, and in allowing labor supply to adjust. The alternative is referred to as the “neoclassical closure rule,” which uses inelastic labor supply and a flexible wage rate to achieve the equilibrium adjustment (Acemoglu, 2009, p. 30-31). In some literature, these two closure rules are referred to as the short-run and long-run labor market closures, respectively. This is a reasonable interpretation, as in the long run one would expect

---

4 An alternative to the constrained approach is to restrict the availability of an input by way of a "phantom tax", which raises the input’s price to a level that limits its demand to what would otherwise be the constrained level. It is referred to as "phantom" because the tax revenue is short-circuited from being spent so as to avoid unduly affecting other aspects of the analysis (see, e.g., Dixon et al., 2010; Giesecke et al., 2012).
that labor mobility and various adjustments would bring about a normal equilibrium. However, the downside is that most applications of the model using this closure rule to examine a proposed policy change will result in no change in employment. While employment is not a welfare measure, it is of significant interest to policymakers, and results of zero employment impacts often raise concerns from this closure rule while being in the spirit of OMB guidelines which assume full employment. Some models fully endogenize labor and so employment changes can result but are not themselves welfare measures.

One might think that the long-run closure rule appropriate to most applications of benefit-cost analysis involving projects and the typical effects of policies intended to be established for long duration. However, the short-run (Keynesian) closure rule would be applicable during the construction phase. Otherwise, the appropriate choice of close rule is an empirical question as to whether labor is fully employed or not.

Most texts on benefit-cost analysis admonish the reader against including general equilibrium or other types of “multiplier” effects, citing that any gains in other markets must come at the expense of other activity because fully employed resources must be diverted. After numerous examples, however, many texts include a statement on the order of: “Local projects are most likely to generate significant positive benefits in secondary markets when local rates of unemployment are high or other local resources are idle.” (Boardman et al., 2011). Of course, this places a burden on the analyst to determine the level of employment, not only in the market in question, but elsewhere in the economy. At the same time, the aforementioned statement and the response to it often neglects the fact that it is less relevant in the case of an analysis performed in context of regional economy, or in a national economy with open or porous borders and otherwise ease of mobility. For example, at the regional level in the US, it is not unreasonable to assume that additional labor will migrate (or commute) into region to fill job openings from neighboring regions where unemployment exists.

Ultimately, GE models explicitly solve for multiple market equilibria perhaps with various Government or other closure constraints on the equilibrium (Florio, 2014; Mas-Colell, Whinston and Green, 1995). One part of that interaction is the presence of terms from “other” markets in
any one particular market equilibrium. A common structure involves the presence of prices of substitutes or complements such that cross-price derivatives (and hence elasticities) exist (Goulder and Williams, 2003; Harberger, 1964).

Chetty (2009) distinguishes between the sufficient statistic approach of Harberger, and structural estimation approaches involving specification of an entire model, essentially the GE approach. The presence of effects from other markets in the sufficient statistic for welfare impacts goes back to Harberger’s (1964) market analysis. To the extent that cross-price elasticities are non-zero and there are distortions in other markets, then analytically there is an impact in the related markets. Goulder and Williams (2003) suggest in particular that the labor market distortion creates a large divergence between PE and GE approaches, although the ultimate test of error and bias is with observational data. Ultimately, whether the GE impacts are large or small depends on the size of the change in the original market, the cross-price derivative, and the size of any distortion in the market and the accuracy of the maintained hypotheses. Restrictions on the absence of other elements becomes important for PE analysis; but for GE analysis, the presence of such parameters model the behavior of consumers such that the entire economy is sensitive to a change in any one market which is then propagated until a new equilibrium is reached. The estimate of any impact is conditional on the maintained hypotheses of the model, such as the equilibrium or other closure restrictions.

3. Relevance to OMB Guidance and Conclusion

The OMB proscription against including secondary or multiplier effects, by its very wording, is based on an earlier generation of GE models, namely input-output (I-O) modeling. In guidance for benefit-cost analysis (OMB, 1992), additional mention is made that at full employment there can be no economy-wide secondary or multiplier effects, presumably in aggregate. The dominant (computable) general equilibrium model in use at the time, and still often used today, was I-O (Rose, 1995). The linear algebra behind the model requires that an increase in activity generates, in the terminology of the model users, additional secondary or multiplier effects. Such a mathematically “locked in” result was thought to be abused in the analysis of various
projects and policies and so a kind of Type I decision error avoidance became enshrined, to exclude such indirect impacts as, in most cases, it would be irrelevant in the aggregate.

Modern GE models are built on solutions to non-linear feedbacks and equilibrium conditions in markets so that positive, “general equilibrium” multipliers are not a required outcome of the model. Depending on the form of labor market closure, constant employment, full equilibrium employment, or sustained unemployment can be modeled. Hence newer GE models address some of the concerns about use of GE models to inform policy decisions.

Ultimately, many policies of interest to government decision-makers are not incremental policies, even if much heat may be generated in such cases. Polices related to the control of green-house gases, homeland security expenditures, health care, international trade agreements or large scale government expenditure programs to expand the economy are not small changes likely to alter prices or constraints significantly in only one or a few markets. For such questions a GE framing seems appropriate. Other policies may require modeling at a level of detail or non-market activities which may be difficult to analyze using a GE approach.

Consequently, given the common theoretical constructs for welfare measurement used in both PE and GE models and the evolution in computation, the proscription against GE models in benefit-cost analysis for regulatory, policy, and program purposes appears unwarranted and worthy of review. This is not to say that either the PE or GE approach is truth, they are abstractions of reality and will continue to evolve.

References


