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

Scott Farrow and Adam Rose

Abstract

Advances in theoretical and computable general equilibrium modeling brought their conceptual foundations in line with standard microeconomic constructs. This reduced the theoretical and empirical gap between welfare measurements using a partial or a general equilibrium approach. However, the separation of the partial and general equilibrium literatures lingers in many applications which this manuscript seeks to bridge. The now shared conceptual foundations, the importance of functional specification, the role of common price movements, and closure rules are discussed. The continuing US Government exclusion of secondary effects from welfare measures in some applications is questioned.

1. Introduction

There are two schools of practice for applied welfare analysis, partial and general equilibrium analysis (PE and GE respectively). While there exist 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 are codified by the US Office of Management and Budget (OMB) for both regulatory analysis and benefit-cost analysis in general. This note seeks to bridge the practitioner’s divide between partial and general equilibrium by reviewing the substantial commonality between PE and GE assumptions, while identifying some of the differentiating complexities that arise in analyzing some issues particularly those related to heterogeneity, distortions, and distributional issues. An additional purpose is to review the basis for excluding secondary or indirect effects in OMB guidance for the benefit-cost analysis (BCA) of government investments, policies and rules. The

---

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 V Kerry Smith for comments and to the Department of Homeland Security for funding support. This version October, 2016.
Within GE analysis, we observe two strains. One is multi-market models in the spirit of pure microeconomics, such as computable general equilibrium (CGE) analysis (Shoven and Whalley, 1992; Dixon and Jorgensen, 2013). These models typically consist of many economic sectors and explore the interactions between them. They typically use the concept of a representative agent (producer or consumer), as the decision-making unit and of the sector. These models may lack some of the financial components of the second strain of theoretical and applied “pure” macroeconometric models, such as the money supply, and the complexities of interest rate determination. We confine our attention to macro models based on micro foundations rather than “pure” macro models although the gap between even the two macro approaches may be diminishing.

Partial equilibrium 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 texts 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. Just, Hueth and Schmitz (2004) in a more advanced text build from a PE approach but cover GE in more detail. The PE approach is also institutionalized in US Government guidance for the analysis of major regulations and other applications using benefit-cost analysis (OMB, 1992; 2003). The guidance precludes consideration of “secondary” or “multiplier” effects (OMB 1992)², although a limited category

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 and proscribed 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)
exists for specific PE elements of “ancillary” benefits or costs unrelated to but caused by the “direct” effects of a regulation (OMB, 2003)\footnote{“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 (2003)}. 

In contrast, a GE approach appears more widely accepted in Europe (Florio, 2014; Dreze and Stern, 1987) and in areas of application that tend to cross many market boundaries such as macroeconomic growth, international trade, taxation, and major terrorism events (Dixon and Jorgenson, 2013). The GE framing typically 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. Some distortions such as environmental externalities have a reasonably long history of inclusion but tend to be the exception rather than the rule (e.g. Hazilla and Kopp, 1990; Kokoski and Smith, 1987). 

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. 

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 sets of models can
provide “shadow prices” for variables of concern, with GE models perhaps better suited to social, trade, or large terrorism issues that tend to involve many markets and challenge PE modeling. At the same time, GE models transmit the smallest of shocks, dampened or amplified as appropriate, throughout the system including effects on input markets such as labor supply. Which model is more liable to ex-post error is little studied compared to studies that assume the GE model is correct. At the same time, the Government’s negation of “multiplier” effects seems to be a historical artifact of a time when simpler GE models were applied to regulatory and expenditure analysis.

2. Welfare Metrics and Assumptions

Applied welfare analysis in the form of benefit-cost analysis seeks to answer the question: when is society’s welfare improved given the investment alternatives under consideration.

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 (Mohring, 1971; Mas-Colell, Whinston and Green, 1995, p. 80-85). Furthermore, the metrics are equivalent when there are no income (wealth) effects as when utility functions are quasi-linear (Mas-Colell, Whinston and Green, 1995, p. 24, 83).

Until 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 and 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 micro-economic foundations aggregating up through individual actors, to markets, to economy-wide analyses and typically with more attention to dynamic processes (Acemoglu, 2009, chapter 5). Prior to this, input-output models, which contain many inherent limitations such as assumptions of perfectly elastic
supply response and absence of market considerations, were prevalently used for empirical multi-market analysis (see below).

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 typically involving a higher degree of aggregation into product classes, ultimately equilibrated and further 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 initial focus and PE will then be defined as restrictions, or caveats, about the assumptions used for the GE analysis while the converse limitations in typical practice of GE are also discussed.

**Figure 1: 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>Sometimes</td>
<td>No</td>
</tr>
<tr>
<td>General Equilibrium</td>
<td>□-----------------</td>
<td>□</td>
<td>Sometimes</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) Marshallian 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). Applied studies often assume no wealth effect or equivalently the quasi-linearity of the utility function and hence the equality of EV, CV and S measures (Mas-Colell, Whinston and Green, 1995, p. 83; Varian, p. 163). Or analysts may be relying on bounds on the estimation error when S is used in place of CV or EV when the (absolute value) of the income elasticity times the S share of income is less than a specified value (Willig, 1976; Just, Hueth and Schmitz, 2004, Section 6.B). To the extent aggregation occurs—with more on the conditions for exact aggregation below—not only individual but aggregate measures of EV, CV and S can be estimated.

Additional metrics are used in GE analysis which 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 almost necessarily omit non-market activities or externalities which are often central to the policy issue at hand. Also, 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).

A change in any of these welfare measures due to an economic shock can be interpreted as a shadow price. 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 whose welfare justification can vary widely as well as conditional welfare measures using value transfer functions to take into account 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 for cross alternative comparison (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). Such adjustments are not common but are certainly present on the frontier.

2.2 Aggregation

Aggregation adds layers of assumptions to the analysis of a single actor, and the PE and GE schools start to diverge. Individual actors and differing commodities are typically aggregated. GE models commonly 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 can there be a more parsimonious representation of the aggregate? Mas-Colell, Whinston and Greene (1995, p. 105-122) break this concern into three parts including when aggregate demand depends just on aggregate wealth, whether aggregate demand carries over all the properties of individual demand, and when welfare measures can be derived from aggregate demand. The latter is the more demanding but of most importance for benefit-cost analysis.

The existence of income (wealth) effects is a confounding factor in aggregation to even 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 to aggregate by sub-group or individuals (Acemoglu, p.
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 a single aggregate, representative 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, however.

But where does concern for a Social Welfare Function enter? 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 occur such 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. 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 and common marginal utility of income for aggregation (consistent with the Gorman form). 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. None-the-less, standard practice for both GE and PE is to aggregate consumers ignoring wealth (income) effects as is done implicitly when S is assumed equal to CV and EV.

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; Just, Hueth and Schmitz, 2004). 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 rationalized through separability restrictions on utility or by the commonality of price movements whether deterministic or subject to a random error (Varian, p. 147-154). 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 scope of economic activity. 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 market distortions such as unpriced or incorrectly priced polluting or non-market behavior such as criminal and recreational choices. Any model is an abstraction, but the change in welfare estimated from any welfare 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 although the omission of non-market factors may be more prevalent in GE models than in PE models which are more frequently designed to address such issues.

2.3 Closure Rules and Aggregation

The third type of aggregation may be thought of as the way in which changes in one market interact with other markets to define an economy wide solution. It goes almost 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 in various ways depending on the functional
form. Such market equilibrium assumptions are standard in both PE and GE analysis. 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). PE models can also, of course, include such restrictions.

The complexity of GE analysis typically requires additional closure rules regarding major account balances in a macro economy in order to solve the system of GE equations. The main consideration is whether one assumes these accounts are in equilibrium or disequilibrium, though this is often couched in terms of which variables are exogenous and which are endogenous (Burfisher, 2011). Major accounts or markets to which this applies include the 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, often explicit in GE models and implicit in PE models. One major approach is often termed the “Keynesian-closure rule”, which allows for an under-employment equilibrium by the device of fixing (holding constant) the wage rate, and in allowing labor supply to adjust (Boardman, et al, 2011). 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 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 will result in no change in employment due to a shock. While employment is not of itself a welfare measure (although it may have welfare implications), it is of significant interest to policymakers. Concluding that there are zero employment impacts often raises concerns while being in the spirit of OMB guidelines which

---

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., 2011; Giesecke et al., 2012).
assume full employment. Some models fully endogenize labor and so employment changes can result.

One might think that the long-run closure rule appropriate to most applications of benefit-cost analysis involves a 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 along the lines 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, this statement may be less relevant in the case of a regional economy, or in a national economy with open or porous borders and 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). Key to the extent of 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). In fact, it is the number of markets to which this assumption is applied that is the major distinction between GE and PE analyses.

To the extent that cross-price elasticities are non-zero then analytically there is an impact in the related markets (Just, Hueth and Schmitz, 2004, p. 346-349). The presence or absence of
distortions in the rest of the economy affects the extent to which equilibrium market adjustments can be assessed solely in a primary market of concern (Just, Hueth and Schmitz, 2004; pp. 327-335). Goulder and Williams (2003) suggest 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 comparing actual and predicted outcomes and not necessarily assuming that a GE model is the true model of the economy. Ultimately, whether GE impacts are large or small depends on the size of the change in the original market, the cross-price derivatives, the size of any distortion in the market and the accuracy of the maintained hypotheses. Restrictions on these elements becomes important for PE analysis; but for GE analysis, the presence of market interactions represents the behavior of consumers such that the entire economy is sensitive to a change in any one market. The estimate of any impact is conditional on the maintained hypotheses of the model, such as the equilibrium or other closure restrictions as well as the data and estimation procedures.

3. Relevance to OMB Guidance and Conclusion

The OMB proscription against including secondary or multiplier effects in benefit-cost anlaysis, by its very wording, was done during an earlier generation of GE models, when input-output (I-O) modeling was the primary tool for computable GE analysis. 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. However, in the I-O general equilibrium model in use at the time, and still often used today (Rose, 1995), the linear algebra behind the model requires that an increase in activity generates 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 (false positive) 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 earlier concerns about use of GE models to inform policy decisions.

Ultimately, many policies of interest to government decision-makers are not incremental policies. Polices related to the control of greenhouse gases, homeland security expenditures, the health care system, international trade agreements or large scale government expenditure programs to expand the economy are not small changes where effects occur in only one or a few markets. For such questions a GE framing seems appropriate. Other policies, including many but not all regulations, may require modeling at a high level of detail or consideration of 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 today and the evolution in computation, the current, default 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, but rather that both are abstractions of reality and will continue to evolve as they seek to improve their analytical power and accuracy.

References


