Journal of Benefit-Cost Analysis

Volume 3, *Issue* 2 | 2012 | *Article* 5

Developing General Equilibrium Benefit Analyses for Social Programs: An Introduction and Example

H. Allen Klaiber, *Ohio State University*
V. Kerry Smith, *Arizona State University, Tempe*

Recommended Citation:

Smith, V. Kerry and Klaiber, H. Allen (2012) "Developing General Equilibrium Benefit Analyses for Social Programs: An Introduction and Example," *Journal of Benefit-Cost Analysis*: Vol. 3: Iss. 2, Article 5.
DOI: 10.1515/2152-2812.1083

Developing General Equilibrium Benefit Analyses for Social Programs: An Introduction and Example

V. Kerry Smith and H. Allen Klaiber

Abstract

Benefit-Cost analyses of social programs often involve complex interactions between market and non-market goods. In this paper we outline one conceptual framework for incorporating the general equilibrium effects into benefit-cost analyses of changes in social programs. The goal is to develop models that are capable of capturing the interactions that take place between decisions involving marketed goods together with their effects on non-market goods and the reverse –changes in non-market goods that cause induce changes in decisions about marketed goods. These types of feedback effects are likely to result from large changes in social programs. These effects are compounded by the potential for endogenous realized outcomes that arise as a result of household actions taken in response to proposed policies. To demonstrate the importance of our conceptual discussion of these general equilibrium effects, we present an empirical example of our general equilibrium framework using teacher cuts in response to budget pressures in Maricopa County, Arizona. We find substantial differences in willingness to pay measures between models which account for endogenous education outcomes in a general equilibrium setting and those which ignore the linkages between market (housing) and non-market (education) goods in a partial equilibrium setting

KEYWORDS: general equilibrium, social programs, education

Author Notes: Partial support for this research was provided by the MacArthur Foundation thru the Evans School at the University of Washington. Thanks are due to Scott Farrow and Rob Williams for careful and constructive reviews of earlier drafts of this paper, Michael Kaminsky for outstanding assistance in developing the data for the analysis of the school performance of Maricopa County public schools and to Jon Valentine and Natalie Cardita for making sense of our initial patchwork draft that resulted in this manuscript.

I. Introduction

This paper describes an approach for incorporating general equilibrium effects into benefit-cost analyses of social programs. To make our description tangible we selected a specific example: the evaluation of reductions in the resources available for public primary education. To highlight the general equilibrium effects of exogenous reductions in the resources used to produce education and its effect on measures of the quality of education, we use a locational sorting model applied to school districts in Maricopa County, AZ, USA. Several of these districts experienced teacher cuts in the 2009–2010 school year. These cuts provide a tangible basis for illustrating how the model would work. General equilibrium effects allow non-market feedbacks to be aligned with the interrelationships between markets that are widely acknowledged in conventional multi-market analyses. In the context of sorting models general equilibrium effects influence our understanding of *both* the severity and distribution of changes in household well-being arising as a result of changes to local social programs.

Most discussions of the distinctions between partial (PE) and general (GE) equilibrium frameworks for benefit-cost analysis focus on policies that directly alter the prices of marketed goods and services. Many follow the seminal contributions described in Just, Hueth, and Schmitz [2004]. When these analyses consider policies altering the amounts, quality or conditions of access to non-market resources, there are added complexities. Nonetheless, the basic concepts defining PE and GE measures of net benefits are comparable.

After providing context for our analysis, section three defines benefit concepts and explains the difference between PE and GE welfare measures. Section four describes two approaches for developing models capable of measuring GE effects with non-market interactions that might arise from social or environmental policies. The first of these approaches uses the computable general equilibrium (CGE) logic originally developed by Scarf [1973] and made easily accessible through the MPSGE framework developed by Rutherford [1997].[1] The second modeling approach involves locational sorting models. The locational equilibrium approaches are different from the CGE logic. They do not represent GE welfare measures in the same way that the conventional CGE framework does.[2] We discuss locational equilibrium models because they are especially

[1] Shoven and Whalley [1992] provide an introduction and early review of this literature.
[2] As we discuss in the next section, the use of the term general equilibrium can have a number of interpretations from considering effects through more than one market to a situation where the model describes a policy's effects on prices, incomes and non-market services affected by market decisions. This is the definition we use when we refer to general equilibrium analyses. The

relevant to situations where social programs are provided by local governments. To extend the CGE framework to include market and non-market policies, the framework must describe how the services that result from these programs influence preferences and/or production activities. These specifications, in turn, influence how non-market feedbacks influence model outcomes. As a result, we include a discussion of how most CGE models have addressed these issues. Our review is somewhat critical of the most popular maintained assumption of separability between market and non-market influences. To help document our argument, we describe some of the existing analyses of PE versus GE measures associated with environmental policies.

Finally, our example of PE versus GE welfare measures for educational policy is developed in two sections. Section five describes how a vertical locational sorting model is estimated to characterize household preferences for the quality of local public education. Our application is based on housing sales data and the differences in performance of primary school students on standardized tests for 46 school districts in Maricopa County, AZ, USA. Our analysis allows school quality to be endogenous to equilibrium household sorting and solves the model as a Nash equilibrium. Each household is assumed to respond to housing prices and to recognize what other households will do in response to exogenous changes in resources for public schools. Households decide to relocate based on both housing prices and the implied effects of all household movements on the resulting school quality in each district, which is influenced by student/teacher ratios.

An outline of the conceptual features of this model and our example describing the PE and GE costs of eliminating different numbers of teachers in each of a subset of the school districts in Maricopa County are presented in section six. The last section summarizes what is known about developing GE benefit-cost assessments for social programs and outlines a few areas where further research would be likely to yield significant payoffs.

II. Background

A. Conventional Practice

Most of the economic analyses of the GE consequences of policies have been in the context of marketed goods and services. The focus has been on the effects of distortions in markets or the effects of new interventions. These can be pre-existing taxes, environmental regulations, or non-competitive markets. These effects are represented as creating wedges between what demanders pay and

---

definition of these measures in each modeling strategy is conditional on the assumed structure of the choices made by economic agents and the rules governing their interactions.

suppliers receive for the goods and services involved. There have been a number of interesting findings. For example, a pre-existing distortion can influence the measured effects of a new tax depending on whether a partial versus general perspective is adopted. The measured excess burden from introducing a new, relatively small tax can vary substantially due to GE effects (see Goulder and Williams [2003]). In another context, Goulder et al. [1999] demonstrate similar GE "surprises" can arise in comparing the cost-effectiveness of environmental policy instruments with pre-existing distortions. These existing distortions raise the efficiency costs of all environmental policies relative to the costs in a first best world.[3] In addition, they can alter the superiority of tradable permits over performance standards and mandated technologies from an efficiency perspective.

The evaluation of the differences of PE and GE effects cannot be distinguished from the assumptions made in defining the "size" of an intervention in relation to one or more markets. In most public economics applications involving taxes or regulations the evaluation typically focuses on the direct effects of the tax or regulation on prices. For these applications, size is interpreted in terms of a policy's effect on a PE or a GE measure of excess burden. The direct effect on excess burden is often proportional to the square of the size of the policy, and the policy is treated as a price wedge. For the case of taxes, size is based on the size of a change in tax rates. In practice, there has not been a specific definition of large or small changes. Early literature such as Shoven and Whalley [1972] or Whalley [1975] motivated their assessments using previous studies of tax distortions. For example, Shoven and Whalley used Harberger's estimates of the efficiency costs of differential tax treatment of capital in the US. Whalley used a specific reform in UK capital tax rates. More recently, Goulder and Williams [2003] used a range of hypothetical new tax rates on commodities in different sectors to illustrate their arguments for the importance of a PE versus GE perspective in measuring excess burden. For these types of examples the GE effects are often approximately linear in this price wedge.

With environmental applications size can have a somewhat different meaning. It can arise from a spatial dimension, such as the amount of undeveloped land that is preserved in a protected status or the number of households affected who might be demanders (or suppliers) of specific goods and services. In either context, size does not correspond to a price wedge in a single market. In these settings, Palmquist's [1992] arguments for how to evaluate

[3] Cost in this discussion is defined as the efficiency cost of a tax. This includes the costs due to abatement, output substitution, and input substitution. When pre-existing distortions are allowed and with them a second best setting is considered there are marginal revenue recycling effects because new taxes can be used to reduce existing tax rates (while maintaining a consistent revenue) and there are interaction effects on measures of the marginal excess burden. See Goulder et al. [1999], pp. 336–337 and pp. 344–346.

localized externalities are frequently used as an analogy to describe how "size" affects benefit measures in differentiated markets. The challenge arises because for environmental applications the change associated with a market response takes place within a specific geographic context. At least two steps are involved in translating it into an equivalent price change experienced by different households.

The first of these steps involves the representation of the spatial features associated with the impacted amenity as an attribute of a location. For example, is open space measured as the amount (acres) within a distance zone of each property, a measure of distance to the closest open space, or both? Depending on the way the non-market attribute is represented, the effects on the hypothesized tradeoffs people would make to obtain (or avoid) it will be different. Equally important, the number of locations affected will be different. These distinctions are important because they imply the selection of a modeling strategy and a metric for measuring the non-market good will influence the "size" of effects assigned to each policy. With this outcome we also have a difference in the potential importance of PE and GE responses. Thus, the characterization of the market and the policy are jointly related in determining the equivalent of a price effect. In this case, it is the price change for one or a small number of locations.[4] Related to this point is the description of how a non-market good is made available to a consumer. Many economic models do not describe the geographic extent of the market. For some applications this is not important. However, it is important for non-market services. Non-market goods have virtual prices that cannot be defined without reference to the geography of the processes that influence their availability.

These issues arise in many forms. Here, attention is focused on the geography for considering the distinction between GE and PE welfare measures. The geographical extent of the market must be assumed for any benefit-cost analysis. The Clean Air Act introduces regulations that improve air quality in the US and may also affect air quality in Canada or Mexico. Whittington and MacRae [1986] considered the issue of who should "count" in benefit-cost analysis as a question of who has "standing." For national policies and associated rules, we may be comfortable with using the political entity responsible for the rules to provide this definition.

However, if we consider a local change for an important resource the question of the extent of the market may not be easily resolved in these terms. Removing dams on the Klamath River affects a river system. It is a federal decision with input from the states and stakeholders but the extent of the market for the changes in environmental services is not as clearly national.

[4] Nonetheless, as Kuminoff and Pope [2010] demonstrate this argument is best interpreted as an approximation.

Two important distinctions arise in these examples. First, what domain defines the affected individuals and do their responses influence markets that might signal tradeoffs they would make to get (or to avoid) changes that are taking place due to an exogenous policy change. Second, are these services also influenced through a natural system that responds to what people do in response to an exogenous policy change? This system might be related to spatial markets but is separate and not market based. It leads to what we label here, non-market feedbacks, and need not have the same spatial definition as the market system used in describing a GE response.

B. Just, Hueth, and Schmitz's General Equilibrium Demand Functions

In a conventional summary of GE analysis in a static multi-market setting, such as Just et al. [2004], the analysis would begin by describing the features of GE demand functions.[5] When an exogenous distortion, such as a tax or a regulation, is introduced into an otherwise undistorted market equilibrium their analysis demonstrates how the PE and GE benefit measures can be compared in a single market. This explanation is based on the concept of a GE demand function. In the GE demand function the measure of how the quantity demanded of a good change with a change in its price is defined so that this response measure is adjusted by the contributions that substitute or complementary markets make to equilibrium outcomes (observed in that good's market because its price has changed). When demand and supply functions describing all markets are linear the derivation is straightforward. The reduced form expressions for equilibrium prices in all other markets are used to eliminate their price effects in the demand for the commodity of interest. The term describing the price effect is the algebraic counterpart to assuming equilibrium adjustments in all markets have taken place when evaluating a price change in the one market of interest. It is consistent with GE welfare measures because there are no distortions in any other market. By assumption, each consumer's marginal willingness to pay is equal to the marginal cost for each good in these other markets. The action in the market of interest does not change that outcome. It may change the levels of consumption of these goods, and correspondingly the levels for their marginal willingness to pay and marginal cost in the new equilibrium. Only the market with the intervention has a wedge if the intervention is a tax or regulation that affects demanders and suppliers differently.

Harberger's [1964] strategy for measuring excess burden exploits a similar logic for the case of a single distortion. Recently, in arguing that the Harberger formula can be interpreted as a sufficient statistic, Chetty [2009] outlines two

[5] See Chapter 9, pp. 311–350. Appendix 9B provides a detailed summary of general equilibrium welfare measurement for changes in a market economy.

conditions. The first is that behavioral responses to price changes for other goods, induced by a new tax on a separate commodity (or service) are not first order effects. In this context, he means that the majority of the effect will be captured in the market experiencing the change. He is making an analogy to a first order (linear) approximation which would ignore "interaction" effects. The second is that in the absence of other distortions, the other price effects cancel between consumers and producers. Of course, this assessment does rely on our ability to measure the full effect of the new tax on the good being taxed.

When we consider the analog to a GE demand for non-market goods a parallel result could be suggested. That is, if a policy intervention affects multiple uses of an environmental resource we might not observe adjustment in response to a change in one context but would observe it in another. This would not mean "equality" of marginal willingness to pay and marginal cost (i.e. price and marginal cost in the context of multiple markets without imperfections). Here, it might mean considering the costs of adjustment along some margins such that the differences would not be large enough to motivate changes. The natural question that might arise concerns the mechanism that would lead to these types of outcomes.

One possibility would be through a hedonic model, but for the question to have meaning we would need to consider a non-market good that served multiple purposes for the same household. In this context, we could observe it affecting different types of choices. The reason is that one might be assumed to be capable of being adjusted by adapting behavior with relatively low cost and another might require larger adjustment costs. The example discussed in Phaneuf et al. [2008] is water quality in local lakes and its influence on housing prices and on local recreation trips. The changes in water quality could be small enough that they did not influence the hedonic equilibrium (e.g. the disparity between marginal willingness to pay and marginal price was not large enough, given the transaction costs of moving, to warrant relocation) but there could be a reason to measure the gains in terms of local recreation.[6] Recreationists would change the location and number of trips based on the change in water quality. Larger changes would be required to see the market feedback through housing markets. As a rule, however, the ways that are available to households to select different amounts of non-market goods are limited.

III. How Should the Net Benefits of a Public Intervention be Defined?

The question posed in this heading is general. Our focus here will be targeted. We accept a conventional economic definition for what constitutes a benefit. It is

[6] We are grateful to a referee for suggesting the link to hedonic equilibria.

either the amount a person would be willing to give up to realize some specific change from his (or her) baseline conditions (in the case of a willingness to pay for a desirable change) or it is the amount of compensation a person would require to be indifferent between having the change versus remaining with the baseline condition and receiving monetary compensation (the willingness to accept measure). A key element in the PE versus GE distinction is the characterization of "the change". In most standard discussions it would be described as a price, quantity, or quality change. However, when one considers policy evaluation, a key question is translating the change we associated with the specific policy, into a change in one or more of the exogenous factors that influence people's (and firms') decisions. This translation is directly connected to the differences we will classify into PE and GE measures of net benefits.

A. Some Definitions

A Hicksian consumer surplus assumes we can represent a person's choices within a constrained utility maximizing framework. In the simplest static example, the constraint is a budget constraint relating income to the prices of the goods and services consumed. The outcome of these choices is represented with an indirect utility function, where realized well-being, *V*, is a function of income, prices, and any other exogenous factors contributing to well-being and outside a person's control.

We will adopt the convention that a superscript zero (0) defines the baseline condition and a one (1) is the new or altered condition with the policy. Equation (1) defines the well-being that a person realizes in the baseline as a function of income, prices, public goods, and other exogenous factors contributing to well-being.[7]

$$V^0 = V(m^0, p^0, Z^0, q^0) \tag{1}$$

where m = income, p = a vector of prices of marketed goods, q = a vector of public goods (produced by local governments) and/or available natural assets such as air quality, and Z = a vector of quasi-fixed private commodities assumed outside an individual's direct control.

Willingness to pay is the income a person would give up to obtain some desirable change. Thus, if we assume one price (in our example the price for the first good) is lower in the new state (superscripted with 1) than in the baseline

[7] Our definitions for consumer surplus do not require that we measure utility or well-being. This characterization is an analytical abstraction. It allows formal definitions for the concept being measured (provided we make a set of assumptions about the functional forms for relationships linking variables describing choices and constraints).

condition so the vector p^1 corresponds to $p^1 = [p_1^{\,1} p_2^{\,0} p_3^{\,0} ... p_k^{\,0}]$ and $p^0 = [p_1^{\,0} p_2^{\,0} p_3^{\,0} ... p_k^{\,0}]$, then the willingness to pay (WTP) for this improvement (from p_1^0 to p_1^1) is defined in Equation (2a) and the willingness to accept (WTA) to forego the same improvement is in Equation (2b).[8]

$$V_1(m^0 - WTP, p^1, Z^0, q^0) = V(m^0, p^0, Z^0, q^0) \tag{2a}$$

$$V(m^0, p^1, Z^0, q^0) = V(m^0 + WTA, p^0, Z^0, q^0) \tag{2b}$$

To use Equations (2a) and (2b) to describe the difference between PE and GE measures of WTP and WTA requires a specification for the source of the change in p_1. This source is often illustrated by supposing a tax is introduced. That is, we might suggest $p_1^{\,1} = p_1^{\,0} + t$ for the representative consumer. It is important to recognize this specification makes an implicit assumption about market conditions by maintaining that there is a perfectly elastic supply for the good being taxed. We can assume consumers are heterogeneous in the ways that they are affected by the tax. These need not be considered, as Chetty explains, because what matters for aggregate welfare change is the total change in behavior induced by the new tax. Assuming this can be measured at the aggregate level, compositional effects would not influence the overall assessment. They would be relevant for estimating distributional effects of a tax change.

Consider a more general case; assume there is a policy designated with the symbol θ. It could be a regulation, a tax, or new information about a product or service. The definition is not limited to situations where only one thing changes.[9]

---

[8] For some specifications of *V*(.) we could have a situation where WTP = WTA. This is not the case for a general specification of preferences. A great deal of effort has been devoted to explaining the size of the difference between the two measures. For example, Willig [1976] described the conditions required for the difference between these two measures of consumer surplus to be small *for price changes.* His conditions are based on the size of the income elasticity and the size of the Marshallian consumer surplus relative to the person's income. Hanemann [1991] demonstrated this intuition does not readily apply to situations involving quantity (or quality) changes where the commodity may not be available at a unit price. In this case, the price flexibility of income contributes to the definition of the Willig bounds and does not have a simple relationship to the income elasticity of demand. This distinction is important because we usually have some intuition about income elasticity of demand, considering necessities versus luxuries. The same intuition is not readily available for the price flexibility which describes how the marginal willingness to pay changes with income. The simplest summary would suggest that the relationship depends on the structure of preferences, especially the availability of substitutes for the commodity affected by a policy and how its contribution to individual well-being changes with income.

[9] Chetty's arguments, cited earlier, relate to cases where the first step in the process, the link between the policy and an impact on something that affects consumers' choices such as the tax

It could be a new rule that includes a way to finance the costs of the rule. Thus, there might be a tax change to pay for the enforcement activities required to assure the rule is implemented. In this case, θ represents a vector of actions.[10] A GE measure would consider the change in θ from θ^0 (the baseline conditions) to θ^1. Evaluating this change *requires a model of how the change influences the variables taken to be exogenous from the individual's perspective*. Thus, with direct effects of θ impacting several prices through some mechanism outside each individual's choice process, we would define the GE measure of WTP as:

$$V(m^0 - WTP^{GE}, p(\theta^1), Z^0, q^0) = V(m^0, p(\theta^0), Z^0, q^0) \quad (3)$$

We could assume Z^0 and q^0 also change but, in this case, the process that causes them to change would need to be specified. Income could also be assumed to change either through the revaluation of endowments, a change in the wage, or both.

The specific features of a GE benefit measure cannot be considered to be completely separate from the model used to estimate them. This conclusion follows from the logical structure of each model. That is, different models capture the elements of the policy delivery system that are outside market exchanges in different ways. In the environmental case, most policies are linked to a spatial delineation of people's choices. In the case of social policy, the process can be spatially delineated if the policy is delivered at a local level or is conditional on the local situation. Equally important, there is increasing recognition of the importance of the heterogeneity of consumer preferences to the full outcomes of policy. Often, people can respond outside of markets in ways that influence the policy outcomes. Suppose, for example, the policy involves the quality of local education at the primary level and that we hypothesize that the education outcomes for all students depend on peer group effects. That is, classes with the same minimum number of high ability students experience better learning outcomes for all students than those classes with more homogeneous students, all with average ability. In these cases, the ways in which models incorporate preference heterogeneity and its effects on locational choices and schooling decisions will influence this dimension of their ability to represent GE effects.

---

and price increase, can be identified. In this example, a structural model is required to make this link.

[10] See Hoehn and Randall [1989] for a discussion of piecemeal policy evaluation as compared to a strategy that defines policy composites. They suggest without composite strategies for defining policies the potential biases in using benefit-cost analyses for individual projects can lead to movements away from efficiency when a sequence of individual choices is compared to the composite of the policies treated as an integrated whole. Their argument can be interpreted as reflecting factors that also contribute to the distinction between partial and general equilibrium benefit-cost analyses.

To help develop this point in the remainder of this section we discuss PE benefit measures at the individual level. Partial equilibrium measures select a subset of the effects of a policy. One of the best examples of how this separation of effects can influence the way models are developed is to consider the organization of most benefit-cost staffs conducting policy evaluation. As a rule, one group does costs and another group does benefits. This classification assumes each part can be done separately from the other. In a PE world, we might be willing to accept this assumption. In a GE world, as we explain below, we should not be satisfied with this division of labor.

To illustrate interplay between the definition of the GE measures for the net benefits of a policy and the model used to compute how GE effects arise, consider the argument developed more completely in Smith and Carbone [2007].[11] To begin it is helpful to define the consumer surplus measures in terms of expenditure functions. Thus, we invert the indirect utility function, Equation (1), solving for the total expenditures needed to realize a given utility level. This function is the Hicksian expenditure function.[12] The definitions for WTP and WTA parallel to Equations (2a) and (2b) using expenditure functions, designated here with the function, e(.), are given in Equations (4a) and (4b).

$$WTP = e(p^0, Z^0, q^0, u^0) - e(p^1, Z^0, q^0, u^0) \tag{4a}$$

$$WTA = e(p^0, Z^0, q^0, u^1) - e(p^1, Z^0, q^0, u^1) \tag{4b}$$

u describes the utility level with the superscript of zero (0) corresponding to the baseline condition and one (1) the new situation. Our definition follows the standard convention in applied welfare economics.[13]

In our example of a tax, we assumed $p_1{}^1 > p_1{}^0$, where the subscript 1 indicates this price is for the first good. We would expect it is necessary to spend more with the higher prices to realize the baseline utility. Similarly, the level of well-being with the higher prices would be lower in the new situation ($u^1 < u^0$) so that income would need to be reduced in the baseline condition to be equivalent.

A second issue of interpretation arises when we consider the connections between Equations (2a) and (2b) with Equations (4a) and (4b). The definitions acknowledge that income (or expenditures) are constant when evaluated at the baseline and at the new price and utility so

---

[11] Their analysis was developed to explain the conditions required for Hazilla and Kopp's [1990] analysis of the social costs of environmental regulations to be valid.

[12] See Diamond and McFadden [1974] for discussion of the duality features of the function and its role in early public economics.

[13] See Freeman [2003], pp. 53–63, for further discussion.

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

$$m^0 = e(p^0, Z^0, q^0, u^0) = e(p^1, Z^0, q^0, u^1) \tag{5}$$

This relationship implies that the two measures could also be defined in alternative ways. They might be described as different ways of evaluating (monetizing) the change in well-being due to the policy (i.e. using the new prices – WTP or using the old prices – WTA).[14]

Now turning to why the organization of policy analysts can be important. Suppose benefits and costs of policies are computed separately by distinct groups. Assume further that the tax policy in our example uses the tax revenues to improve q from q^0 to q^1. The WTP measure for an individual's benefits, ignoring effects of the change in p_1 and q on other prices, would be given in Equation (6).

$$WTP^{PE} = e(p^0, Z^0, q^0, u^0) - e(p^1, Z^0, q^1, u^0) \tag{6}$$

Here the sign of WTP^{PE} is not clear. Indeed, this might be described as the individual level net benefits of introducing the new tax, t, to improve q from q^0 to q^1. Hazilla and Kopp [1990] used this logic to define the "social costs" of policy. Their example was environmental policy. They asked about the importance of a GE perspective for measuring these social costs. Before considering this point, we might ask what conditions are required for Equation (6) to be consistent with separately measuring the costs and benefits of a policy. One answer, as Smith and Carbone [2007] note, follows from the definition of the expenditure function, provided it is separable in q, and can be written as:

$$e(p^0, Z^0, q^0, u^0) = \tilde{e}(p^0, Z^0, u^0) - h(q^0) \tag{7}$$

In this case, the net benefit measure can be separated into a cost and a benefit computation, as in Equation (8).

$$WTP^{PE} = \underbrace{\tilde{e}(p^0, Z^0, u^0) - \tilde{e}(p^1, Z^0, u^0)}_{\text{Incremental cost}} + \underbrace{(h(q^1) - h(q^0))}_{\text{Incremental benefit}} \tag{8}$$

With this restriction, the distinction between GE and PE welfare measures depends on whether the analysis takes account of the full price effects of a policy change.

That is, we assume there is a policy change from θ^0 to θ^1 and we ask how, given everything else is held constant, does this change influence the set of

[14] This characterization – monetizing utility changes – has caused considerable confusion in the literature, as summarized by Freeman [2003].

exogenous conditions that constrain each individual's ability to realize a given level of utility. The policy could be confined to one market, or a small number of markets. Alternatively, it could be assumed to influence conditions outside markets. The PE/GE distinction arises in how we convert the policy changes into changes in the prices and other exogenous conditions that influence people's choices. Thus, this is one of the ways there can be an interplay between the modeling assumptions used to describe how a policy influences market outcomes, the assumptions used to define preferences and the resulting PE/GE distinctions.

In the context of education programs, Heckman [2006, 2008] has argued that efforts to improve preschool programs for children will improve the educational outcomes associated with these children. He notes this improvement is also associated with lower rates of teen pregnancy, drug addiction, crime, and other consequences of high dropout rates. It also limits the choice sets these children may experience in their future activities. These added choices can involve further education or jobs for some teenagers and young adults. Do we count all of these effects as a result of the policy improving preschool programs or as just a subset? This question illustrates the GE versus PE question. The issue of model interplay arises here because the decision of what is counted often depends on the reliability of the modeling frameworks used to describe the connections between the policies and each of these outcomes. Can the results leading to changes in education outcomes, teenage pregnancy, drug addiction, and so forth be attributed to the policy intervention in a convincing way that leads to reliable predictions? Equally important, can these outcomes be incorporated in an economic framework that describes the processes determining prices and incomes?

In a different context involving education, the side effects do not have to be positive. Improvements in educational outcomes in one school district may well attract more families. This in-migration could lead to greater competition for homes in the neighborhoods that are assigned to the district experiencing the improvement. With a fixed supply of houses in the districts with better schools, we would expect that home prices would increase. Those households owning homes in the district would gain, whereas renters would lose as their rents increase. Indeed, poorer households could well lose because higher rents might force them to leave and accept a lower quality school district in order to obtain affordable housing. These outcomes would not be intentional and might be missed with a framework that failed to allow for the full scope of GE adjustments. They are part of the distributional effects of a policy.

The importance of a GE/PE distinction depends on the size of the program being evaluated, the size of the markets, the social/economic context being described as relevant to the policy, and the assumed interconnections between the market and non-market influences to individual behavior. If a model focuses on the pure exchange of goods and services, the concept of size has a different meaning. Add a small amount of spatial context and ask how well an addition to the tax on gasoline in one county in the US affects the world market price of gasoline and the answer is direct – not at all. However, if we recognize that even within a small region there is variation in the price of gasoline, the effect of such a tax on the prices of other fuels in *that* region may well depend on the volume of gasoline sold in the county relative to adjoining counties that might be part of the same metropolitan area.

Gauging the difference in a PE versus a GE assessment of the welfare effects of the tax depends on what we assume. As Goulder and Williams demonstrate, with pre-existing labor taxes (at 40%) small taxes on a product such as cigarettes can lead to *large percentage* errors in PE compared to GE measures of excess burden. Their analysis does not consider using the tax revenue for a non-market intervention. At a national scale this can make a large difference, as Carbone and Smith [2008] demonstrate using their model of energy taxes. A small tax (5%) with the same pre-existing labor tax and recognition of the non-market effects (on air pollution) of the tax can increase the differences in PE and GE by 100% when leisure and air pollution are complements as compared to *reducing* the discrepancy by approximately 50% when they are substitutes. Of course, the absolute size of the error in excess burden is also large in this case.

Now, if we return to the spatial dimension of the policy – a small regional clinic – then it is unlikely the feedback effects would be important consider at a national level. However, an analysis at a local level might need to consider another type of feedback. If the improvements in the clinic's quality of care were dramatic and capacity at the clinic remained limited, the net results for local patients might not be completely positive. There could well be congestion and delayed treatment for patients due to that congestion.

These examples suggest that market and non-market effects can both matter. Moreover, the spatial scale that we use in representing the non-market process and who is affected by a policy also matters. This is how the choice of models influences the characterization of PE and GE welfare measures. Some models will not be able to easily reflect the spatial dimensions of the social process or the heterogeneity in agents and their opportunities to adjust. Decisions to use simple, market-oriented, GE models to evaluate social processes then "build in" these potentially important conditioning assumptions.[15] Thus, the definitions of GE and PE are straightforward, but the real "action" arises in the definition of the models and their implementation.

[15] These are part of what Chetty assumed were second order effects for a national level analysis. Our discussion does not overturn his judgment for the case of taxes. It does highlight that the situation can be much more nuanced for other policies.

With that background, PE measurement selects a subset of the possible effects of the change in θ and measures WTP for the change in θ recognizing only that subset of changes. One might suggest that GE includes everything, but this statement needs to be qualified. GE includes everything *that is in the model* used to assess the effect of the policy. Here again, we have an example of the interplay between the modeling choices and the interpretations of the PE/GE benefit measures. Examples using an abstract representation of the issues are in Equations (9a) and (9b), with two prices assumed to be affected for the PE case and everything, *p*, *q*, and *Z* for the GE.

$$WTP^{PE} = e(p_1^{\,0}, Z^0, q^0, u^0) - e(p_1(\theta^1), p_2(\theta^1), \ldots p_k(\theta^0), Z^0, q^0, u^0) \tag{9a}$$

$$WTP^{GE} = e(p^0, Z^0, q^0, u^0) - e(p(\theta^1), Z(\theta^1), q(\theta^1), u^0) \tag{9b}$$

Another important issue that is implicit in these definitions and is also described in some of our examples arises from the differences in the WTP measures for different individuals. As previously discussed, when people are assumed to be heterogeneous the PE/GE distinction must be considered together with the distribution of effects across people. There are at least two aspects of this issue.

People may well experience the *same* price changes but react (due to preference heterogeneity) quite differently. Moreover, once we admit effects outside markets, there may not be complete mechanisms to adjust to differences. It is also possible that the adjustment could well have feedback effects on the policy intervention. Households in different circumstances may well be more or less prone to being impacted by these non-market effects as well.

B. Approximations

As a practical matter it is often difficult to measure the full GE willingness to pay. With revealed preference information we usually have the ability to estimate a subset of the demand functions or will specify a choice function that assumes each person has a limited range of decisions and other choices are not affected by the one being studied. Most policy applications of benefit-cost analysis rely on the existing literature to adapt a point estimate to evaluate the benefits from a policy.[16] Here, we will define two of many such approximations and describe one potential GE/PE distinction. These measures use what is sometimes labeled as the virtual price or marginal willingness to pay for a change in something available

[16] This practice is called benefits transfer. See Boyle et al. [2010] for a current review and appraisal of these practices.

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

outside the market. The marginal willingness to pay for a change in one element in q (say q_1) would then be given in Equation (10) using the indirect utility function in Equation (1).[17]

$$MWTPq_1 = \frac{V_{q_1}}{V_m} \tag{10}$$

The distinction between PE and GE measures in this context might arise with whether the virtual price is evaluated at the baseline levels of prices (and other "parameters" entering the indirect utility function) or at the new level, after the policy change. Thus Equation (11a) would be PE and Equation (11b) GE with the superscript (0) designating the point of evaluation.

$$\Delta B^{PE} = \frac{V_{q_1}{}^{0}}{V_m{}^{0}} \bullet \Delta q_1 \tag{11a}$$

$$\Delta B^{GE} = \frac{V_{q_1}{}^{1}}{V_m{}^{1}} \bullet \Delta q_1 \tag{11b}$$

Smith and Carbone discuss these measures as adjustments to measures of excess burden of a tax to reduce pollution externalities. Using a small-scale computable GE model, they find that the differences due to PE versus GE evaluation points for the virtual price can be important. The size of the effect they illustrate depends on how the commodity that is taxed ultimately affects the non-market good of interest. Their results suggest that the role q plays in preferences is also important to the GE/PE effects on these approximations.[18]

---

[17] For small changes this would equal $\frac{\partial e}{\partial q}$; the definition in Equation (10) holds income constant and in terms of the expenditure function, well-being or utility is assumed held constant. If the change is small they are usually assumed equal. At best, we usually have point estimates for marginal willingness to pay. The distinction between Marshallian measures as defined in Equation (10) and Hicksian (based on the expenditure function) arise when we use these estimates to recover the marginal willingness to pay function. We do not consider these issues here – see Freeman [2003] or Bockstael and McConnell [2007] for discussion of the challenges involved with each.

[18] In that example, one good had a pre-existing tax and the presence of it influenced the size of the excess burden for a new tax. The relationship of q to the good with the pre-existing tax was especially important.

C. General Equilibrium Market Demand

As we noted in situations with undistorted markets, the welfare costs (or gains) associated with a change can be evaluated using the affected market alone. They must be evaluated using the aggregates of compensated demand functions across individuals consuming the good involved and aggregate commodity supply functions. Both must be evaluated at the values of the general equilibrium prices after the change.

This ability to focus on one market relies on two key assumptions. First, it assumes all markets are in equilibrium before and after the intervention, and second all the GE effects arise exclusively through markets. As there is only the one market affected, the GE price effects of the intervention on other goods' prices will be the same for households and firms. Thus, the line integral (or equivalently the measure used for the contribution to economic surplus) will be exactly zero, given the equilibrium condition.[19] In the market experiencing the change in price, the size of the price change realized by consumers compared to what is experienced by suppliers will be different. It depends on the relative elasticities of the demand and supply functions. Thus, even though this market is also in equilibrium after the intervention, the sharing of the price change due to the new distortion is different. As a result, we have the consumer and producer contributions to economic surplus.[20]

In the context of non-market effects, the same result would hold in marginal willingness to pay for each non-market good equal to marginal social cost before and after a change. However, aside from specialized cases such as those discussed in our hedonic example there is usually no mechanism that assures that this condition will be satisfied. This typical set of conditions is what motivates the need for a structural model to describe GE responses. It allows analysts to compute the full GE effects and to consider the distribution of their impacts.

IV. General Equilibrium Models with Non-Market Interactions

To this point the discussion has sought to persuade readers that the "action" in GE benefit-cost analysis is in the selection of the models used to describe how policies create GE effects. The theory concerning what should be measured is reasonably straightforward. This section considers two different, static approaches describing a composite of market and non-market GE outcomes. The first extends the logic of computable GE models to include non-market effects. The primary non-market effects that have been represented in these models to date have

[19] This is what Chetty refers to as implication of the envelope condition.
[20] See Just et al. [2004], pp. 360–361, especially Equation (9.49) and the discussion.

focused on environmental externalities. While other social programs could, in principle, be represented there are significant information gaps concerning just how they would be introduced into the models that would need to be addressed. The second modeling strategy is not a full GE framework. It relies on locational sorting models. They describe situations where price determination is limited to multiple land/housing markets. This set of prices can influence and be influenced by non-market outcomes, but the model does not provide a full description of other markets with effects on both prices and components of income, as has been done in the earlier CGE frameworks.

A. Non-market Effects in a CGE Framework

A CGE model is a consistent description of individual and firm behavior that recognizes the joint determination of product and factor market prices. In a static competitive setting, these models assure the conditions for budget constrained utility maximization by individuals and profit maximization of firms yield a price vector consistent with: equilibrium in all markets, budget exhaustion, and zero profits in all sectors. In the simplest cases, production functions are restricted to constant returns to scale and preferences are often assumed to be homogeneous of degree one. Both relationships are often described with constant elasticity of substitution (CES) functions.

CGE models are calibrated to match the observed baseline expenditures on final goods, payments to factors, taxes and government spending (if included in the model), as well as expenditures on any intermediate goods in a base year. Developing the calibrated version of such a model that realizes the consistency condition is greatly facilitated by using the homogeneous CES specification for preference and production functions. Rutherford [1997] has documented how the process of establishing a consistent link is straightforward for this case. In the case of marketed goods for the benchmark case, all prices are normalized to unity. As a result, quantities of the marketed goods can be measured by observed expenditures. Because the objective of the analysis is to evaluate how relative prices change, this normalization does not compromise the models' relevance (given the assumption of homogeneity of degree one for preference and production functions).[21] As a result, after selecting substitution elasticities, the benchmark data can be assembled into a consistent social accounting matrix and a limited number of parameters calibrated to match the labor/leisure conditions.[22]

[21] Rutherford [2002] has demonstrated simple calibration strategies for the share parameters in CES functions.
[22] One of the few stumbling blocks in these models is the calibration of leisure and work time because time allocations that are not associated with hours worked are incompletely measured (see

Non-Market Calibration

When non-market resources are introduced into CGE models, they are either public goods or activities that reduce externalities. In the case of externalities, we might assume that one set of agents in the model produces them and the consumer receives them. However, the price "paid" for the amount received is usually not a per unit price. There may be no price. For the case of public goods, the costs may be covered through taxes on other goods.[23] The consumer may not have a mechanism to choose how much is received.

Considering the task of calibrating the model with these goods, three important issues concern how we measure: (a) the amounts of these goods, (b) the nature of the relationships between them and the private goods or (factor inputs) entering preferences (production functions), and (c) the information we have about the tradeoffs people would make to change the amounts of the non-market goods. Non-separability, together with the fixed level of these non-market goods *from the perspective of the individual*, implies that the preference (production) functions with non-market goods are non-homothetic.[24] This change alters all of Rutherford's conditions for calibrating the GE to match baseline conditions. They no longer hold. Calibration must match another set of conditions that define the source for tradeoff information while producing the social accounting matrix and equilibrium prices in the other goods and services in the model.

Calibration of these models exploits one or more envelope conditions to define a point where each non-homothetic function "appears" homothetic. Starting with a CES specification (with one or more arguments fixed), the same Rutherford calibration strategy used for the homogeneous case can be applied in the non-homothetic provided it is assumed to hold for the levels of consumption, input use, and values for the non-market good(s) consistent with satisfying the envelope condition(s) associated with the virtual prices for the non-market goods. This assumption means we are nesting a GE problem within the calibration.[25] In

---

Fullerton et al. [1984] for a description of the practice using Hicksian and Marshallian labor supply elasticities to resolve this issue).

[23] Epple, Romer, and Sieg [2001] discuss the prospects of developing models where voting determines the level of public goods and a budget balance condition, together with a specified production function for the public goods assures budget balance. This type of structure has also been included in models by Nechyba [1999, 2000] and Ferreyra [2007, 2009].

[24] See Carbone [2005, 2007] for further discussion and Carbone and Smith [2008] for implementation of a model with these features.

[25] Perroni's [1992] demonstration that a homothetic function can be used to represent non-homothetic responses provides the underlying logic for the calibration. This logic is used together with a more general insight developed by Carbone [2005] that is used in several applications. The insight is associated with combining conditions for GE calibration in the Carbone-Smith case or solving hybrid models. For example, in Carbone et al. [2009], a game theoretic model describing a sub-game perfect Nash equilibrium in the permit market for CO_2 permits is combined with a

the case where non-market goods influence consumers, these conditions also require we define virtual income as a sum of expenditures on market goods plus virtual expenditures on non-market goods. The later correspond to what is implied by the virtual prices derived from the envelope condition linking the levels of the fixed (non-market) goods to the other goods consumed for the benchmark solution. Parameters are set so that this amount of the non-market goods, given the externally observed marginal willingness to pay, would be selected to match the level observed in the benchmark case. The calibration also needs to assure the benchmark reproduces the social accounting matrix for market goods, and any relationship defining how the amount of that non-market good is derived. These equations would be conditional to assumptions about substitution elasticities. Once this is solved then the model no longer treats the non-market good as a choice argument. It is treated as quasi-fixed – changing in response to other choices but not through any optimizing rule associated with households in the model. With non-separable preferences the level for the non-market good affects choices of market goods and the levels of these market goods can in turn affect the non-market good, but there is no direct choice of the non-market good. It is determined as a byproduct of other choices.

Examples

There is very little experience with introducing non-market goods within CGE models when they are treated as making non-separable contributions to preferences or production. Espinosa and Smith [1995] appear to offer the first treatment in a CGE framework. They resolved the calibration by assuming the non-market good had a private good serving as a perfect substitute. They also assumed a Stone Geary preference function. These two assumptions allowed them to adjust the translating (or subsistence) parameters so the level of each non-market good in the benchmark was consistent with its "production activities" and with external estimates for the marginal willingness to pay for each non-market good.

---

competitive CGE model to describe trade in marketed goods. The logic relies on describing GE as a mixed complementarity problem (see Böhringer and Rutherford [2008] as one example) and combining it with equations that recognize the strategic decisions or permits are functions of local derivatives which are defined numerically. This allows the GE solution for the vector of equilibrium prices and quantities and the adjacent perturbed solutions for these variables to be solved simultaneously with the permit allocations treated as endogenous variables that result from the definition for the Nash equilibrium. A comparable process allows the definitions for the non-market valuation measures whether virtual prices (marginal willingness to pay) or willingness to pay to be solved simultaneously with the GE and calibrate a non-homothetic model for a baseline that uses properties of homothetic functions. (See online documentation for Carbone et al. [2009] for further discussion.)

De Mooij [2000] used a log-linear GE model (following the Jones [1965] format) to describe how non-separable externalities affect the impacts of an energy tax on employment, pollution, and income. He does not discuss how the non-market component is calibrated and does not consider the consistency conditions required to link market and non-market sectors. Although the welfare effects of taxes as measured with marginal excess burden are discussed for the separable case, they are not treated with non-separable effects.

Carbone and Smith [2008] appear to be the first to treat these issues in general terms. Their findings for a small CGE model lead to three direct conclusions:

(1) Even in cases where the non-market good has a small fraction of virtual income, the substitution or complementarity relationship between this good and a private good (leisure) has a large effect on the costs attributed to a new tax in an economy with pre-existing distortions.

(2) The nature of the production relationship linking activities with private goods to the level of a non-market good is important for the measured effect of this good on measures of the welfare cost of a tax. A comparison of different response functions holding the share of the non-market resource in virtual income constant found large differences in the effects of the non-market sector on GE measures of excess burden.

(3) In a new analysis with several non-market goods, Carbone and Smith [2010] found that GE feedback effects may appear small based on substitution relationships to market goods. However, this criterion may be misleading. Feedback effects can have larger effects in the interactions among the non-market goods themselves. In this exercise, changes in the virtual price of one of the non-market goods was the best indicator of the importance of the changes in feedbacks arising from different substitution or complementarity relationships between *another* non-market good and labor.[26]

Using CGE for Policy Evaluation with Non-Market Goods

Several tentative lessons emerge from the work with CGE models that are relevant to using this strategy for evaluating the GE effects of social policies.

[26] In their example, changes in marginal willingness to pay for habitat provided a better indicator of how the substitution relationship between recreational fishing and leisure influenced the net of benefits of pollution policies intended to improve both. This measure was superior to using a simple comparison of PE versus GE measures of the net benefits of the pollution policy.

First, the structure of the interconnections between the non-market and market goods is important. Even for situations where the share of (virtual) national income attributed to the non-market sector is small, feedback effects can be very important to the difference between PE and GE welfare measures. Second, the strategy used to evaluate the nature of consumers' preferences for these goods also seems to be important. To date, the evidence has been largely through the interaction of estimates of consumers' preferences and the shape of the "production relationship" for the non-market good (i.e. in the case of air pollution it would be the dose-response function). Little has been done to systematically study the affect of alternative specifications for these relationships.

Third, most of these models use simple characterizations of the consumers and of the jurisdictions within which they select private goods. Carbone and Smith assumed a single aggregate consumer and abstracted from jurisdictions providing public goods. For many social programs, there are multiple opportunities for people to adjust by changing locations.

Finally, benefit-cost analyses are increasingly being expected to include measures of the uncertainty in estimates. In the case of the CGE model, there has been limited work on how uncertainty in estimates used to calibrate these models would affect their results.[27] We should draw a distinction between treating the parameters used in GE policy analyses as random variables and sensitivity analysis. The former attempts to develop distributions for the computed GE effects – whether price changes or measures of willingness to pay. The latter recognizes a range of values for potential point estimates and considers the sensitivity of results to different values. It is not clear which strategy is most informative.

There are a number of complex issues in sorting through the arguments associated with estimation uncertainty and economic models. One line of argument might be the estimation uncertainty in our measures of parameters for any model, PE or GE is so large that the best an analyst can do is keep models simple until we can narrow the degree of uncertainty. Another argument would be to argue that one reason for the imprecision we observe is the GE nature of the world. Under this view, we need to do a better job of representing it at all stages – estimation and policy analysis. Another position might be to suggest that all policy evaluations rely on calibrated models so estimation uncertainty is only important if it affects the confidence of those using the models to inform policy judgments. Under this view, full blown efforts to derive the probability distributions induced in PE or GE effects due to distribution for estimates are not likely to be useful.

[27] A notable exception is Harrison and Vinod [1992] who evaluate the sensitivity of CGE results to parameter values used in calibration in a statistical framework.

Our view is that selecting one of these perspectives for all problems does not make sense. At this stage in our understanding of how to use models to inform policy we believe a strategy that relies exclusively on simple assessments is likely to be best. The choice of the type of modeling strategy to adopt should instead be guided by answering some progressively refined questions:

- Is there agreement on the nature and empirical importance of the primary features and parameters of the models needed for an assessment, at least for first order problems, in the literature? If yes, proceed to a detailed model of the system. If no – keep it simple!
- Within a detailed model, can we agree on some stylized facts or parameters that we have confidence have credible estimates within a relatively narrow range? If no, then a large meta-like sensitivity analysis of modeling outcomes is required (see Banzhaf and Smith [2007] as an example). If yes, then a target sensitivity analysis using the parameter(s) with largest variability in estimates is likely to be informative.[28]
- Within a targeted analysis, are there thresholds that identify decision points such as positive net benefits that can be described and corresponding questions about the importance of what has been left out? If no, then the analysis may help to reference the policy choices. If yes, then the GE/PE distinction and structured sensitivity analysis may be especially important.[29]

B. Non-market Effects in a Locational Equilibrium Framework

Some of the services provided by local governments as well as environmental services available at different geographic locations have the characteristics of a local public good. The amounts available vary between the different locations. In the case of environmental goods these differences could be due to variations in natural conditions. For other goods they could be the result of decisions made by local communities. They are what led Tiebout [1956] to suggest that for local

---

[28] Farrow [1999] provides another illustration of an effort to evaluate how elasticities derived from an analytical GE model varied with estimation uncertainty. Using an analytical model developed by Fullerton and Metcalf [1997], he assumed triangular distributions to characterize parameter uncertainty for the marginal labor tax rate, the aggregate income elasticity and the elasticity of substitution. Other variables were represented with uniform distribution and the analytical expressions derived in Fullerton and Metcalf were used to simulate the joint impact of uncertainty in all the parameters on the implied welfare effects of taxes.

[29] See, for example, Goulder et al. [1999], Parry, Williams and Goulder [1999], and West and Williams [2004].

public goods, communities were the "supermarket" allowing households to select the best match, given their preferences and abilities to pay.

A number of authors, in public and urban economics, have attempted to evaluate the analytical properties of Tiebout models.[30] This literature offers two approaches for discussing PE versus GE benefit measures. The first uses calibrated models such as Fernandez and Rogerson [1998] and the second directly estimates a sorting model, and is due to Epple and Sieg [1999]. The Epple-Sieg paper demonstrates how a consistent description of a locational equilibrium could be used to estimate a model that: (a) recovered measures of heterogeneous households' preferences for housing and local public goods, and (b) offered the means to describe how that equilibrium would change with exogenous changes affecting the amounts of those public goods.[31]

The models can be grouped depending upon how they characterize the heterogeneity in household preferences. The first specification has been labeled a pure characteristics model and the second is labeled a random utility model. The Epple-Sieg version or pure characteristics model assumes each household (or each type of household) has a different taste parameter for the locationally differentiated good. However, the households all evaluate the locational attributes or services contributing to the amount of the local public good in the same way. The alternative random utility model allows the parameters of location-specific goods to vary more generally.[32] Our example uses the pure characteristics framework. As a result, we focus the remaining discussion on estimating, interpreting, and using these models for policy evaluation.

Structure and Estimation of the Pure Characteristics Model

A key assumption allowing these models to estimate the locational equilibrium in a way that facilitates estimation is the single crossing condition. To appreciate what this condition implies for preferences we need to develop its structure more specifically. Suppose we can define the price for a homogeneous unit of housing in each location and measure the amount of the non-market good a household acquires by locating in a specific community. The Epple-Sieg model maintains that two sources of household heterogeneity influence their sorting across communities. One is household income and the second is the household taste for

---

[30] A review of these studies is beyond the scope here. See Epple and Platt [1998], Epple, Gordon, and Sieg [2009], Klaiber and Smith [2011b], and Kuminoff, Smith and Timmins [2010].

[31] Several of the calibrated models have considered educational policy but the focus was primarily on accounting for the effects of GE adjustment or outcome measures rather than benefit measures. See Kuminoff, Smith and Timmins [2010] for further discussion.

[32] See Klaiber and Phaneuf [2010] for an example of a horizontal model that develops GE and PE welfare measures. Bayer et al. [2007] also estimate a horizontal model but do not compute welfare measures or consider the PE/GE distinction.

public goods, represented by an unknown (to the analyst) parameter β. Income and the taste parameter are assumed jointly distributed to represent household heterogeneity. To characterize the equilibrium we must assume that both income and the taste parameter satisfy the single crossing condition.

The single crossing condition restricts how the price equivalent change in a good related to the non-market good changes with income and with the taste parameter. We describe it as restricting the price equivalent because the first step in understanding the condition is to characterize a price change in the related market good that is equivalent to a change in the non-market good (in terms of its effects on household utility). Once this price change is defined we ask how would it change with either income or household tastes. This relationship is formally described using the indirect utility function in Equation (1). Assume p_1 is the annual rental price for a unit of homogeneous housing in community one and denote the amount of the non-market good in that community as q_1, then the single crossing condition in income can be defined in Equation (12).

$$\frac{\partial}{\partial m}\left(\frac{Vq_1}{Vp_1}\right) = \frac{\partial}{\partial m}\left(\frac{dp_1}{dq_1}\right) > 0 \qquad (12)$$

The variation in preferences stemming from the taste parameter (β) also must be consistent with the single crossing condition (i.e. $\frac{\partial}{\partial \beta}\left(\frac{Vq_1}{Vp_1}\right) > 0$) for all communities. Taken together these features assure that the sorting equilibrium satisfies three conditions. These three features are especially important to using the equilibrium for estimation and policy evaluation. They include:

(1) When communities are ranked by the public good index and by the equilibrium community-specific housing price and the two rankings are the same (this is called the ascending bundles condition).

(2) With a continuous array of different types of households, there is a set of households indifferent between communities with adjoining ranks based on price and the index of public goods (this is called the boundary indifference condition).

(3) When households are ordered among communities, given a level of income, the equilibrium ordering implies that households with the greatest preference for the locational public good locate in the community with the largest amount of that good. Thus, conditional on income, households are stratified by their taste for the public good (this is called the stratification condition).

Using a CES specification to describe preferences, the three characteristics of the equilibrium allow the ordering of communities by price and the ordering of households by the unobservable taste parameter. This ordering is used to "predict" a distribution of income for each location. This feature provides the intuition for an estimator to recover preference parameters (see Epple and Sieg [1999] for details).[33] An important feature of the model arises from the boundary indifference condition. This property of the equilibrium allows the definition of a relationship between the equilibrium prices for adjoining communities (in the ranking implied by the equilibrium) and the index for the levels of the community-specific public goods. That is, given estimates of the equilibrium price indexes it is used to define recursively an index for the quantity of the community-provided public goods. For policy analyses of how to introduce an exogenous change in one or more community's public goods, this relationship is reversed and used to define how prices "must" relate to the revised indexes of public goods across communities. These prices are used in a framework that derives housing demands, compares demand in each community to supply, and then updates prices and re-sorts households until housing demand equals supply in all communities and no households have an incentive to move given the quantities of public goods and the housing prices in each community.

Examples

There are two published examples of the vertical model that discuss welfare measurement. The first by Sieg et al. [2004] uses the model to describe sorting among school districts in Southern California for educational quality and air quality. They demonstrate how the sorting framework can be used to estimate preference parameters and to compute the benefits for improvements in air quality in a framework that allows a comparison of partial and GE WTP measures. These measures were defined for an exogenous improvement in air quality.[34] The simulated results permit an assessment of the role of preference heterogeneity for measures of the differences in welfare across communities and income groups. This approach also allows the differences in PE versus GE measures to be considered using these same categories.

In their application the difference between PE and GE results are defined by how the change in air quality is treated and whether the housing price effect is

[33] Klaiber and Smith [2011b] provide more discussion of the intermediate steps in the logic in comparison to the horizontal or RUM formulation for the model.

[34] The actual simulation of the equilibrium is numerical. It draws two random variables to characterize each household (i.e. income and the taste parameter for the index of public goods in each community) and selects a large number of these pairs of random variables. Each is assumed to represent a "household".

taken into account. That is, each simulated household "controls" the air quality it ultimately experiences by moving. To implement Equations (2a) and (2b) we propose a thought experiment. For the PE measure we evaluate how the "simulated" households assigned to a location (based on their income and taste for a public good evaluate the housing price and public goods) would evaluate an air quality improvement in that location, assuming housing prices do not change. The GE measure allows households to re-sort, with prices adjusting until the quantity of housing demanded in each community equals supply at the new set of prices and households have no incentive to move. This measure compares the baseline prices and public goods to the new prices and public goods at the locations each household selects under the new conditions. Both the realized air quality through adjustment and the price change contribute to the distinction between PE and GE measures in this case.

Several possible scenarios were considered in this study and the illustrative computations presented in the paper suggest that even though air quality might improve everywhere, some households might experience losses because housing prices increase more than the improved air quality was worth to them. To some extent this result may be due to the simulation's inability to capture exit. The set of communities is a closed system in the simulation. Thus, a household must sort within the locations described by the model.

The second application by Walsh [2007] is similar in structure. The locational amenity in this case is open space and the sorting process is similar to what was used in the air quality application. There is, however, one important difference. Open space in Walsh's model is defined so that it is produced in part through the outcomes of household sorting. In the Sieg et al. [2004] paper, each household can select an air quality from a finite set so there is some endogeneity as part of the equilibrium amount realized by specific households. However, these choices do not feedback and alter the amounts of air quality available in each community. They affect the prices of homes. For the Walsh model, choices affect both prices and the amount of open space! Open space is assumed to be a composite of protected land and undeveloped land. As a result if policies to protect land attract households, it is possible for the equilibrium to lead to smaller amounts of open space by developing more of the unprotected land. In this situation, both the prices of housing and the amount of the local public goods in each community are equilibrium outcomes.

Walsh's research illustrates how a Nash equilibrium can be computed for endogenous determination of prices and open space. Our example in the following section builds on this work and allows the level of the location specific good to be determined jointly with prices through the sorting process.

V. General Equilibrium Benefit Measures for Social Policy: Local Education

In this section, we outline a new application of the vertical sorting model to illustrate the decisions that need to be made in developing a GE analysis of the benefits of a social policy. We consider local education in Maricopa County, AZ, USA. This section and the following section describe how the model was estimated and how it can be used to evaluate the PE and GE benefits associated with changes in the number of teachers available in this county's school districts. In this section, we outline the model, describe the data, and show how the data are used to estimate the model's parameters. Section VI describes a policy analysis with the model.

Before getting started with our example, it is important to emphasize that our objective is illustrative. The model builds on existing research but is not intended as an alternative model for evaluating the gains for policy changes that affect local education. This is an important perspective because there is a large body of literature confirming the importance of school quality on housing prices. The magnitude of the effect is often debated, as illustrated by the contrast between Black's [1999] estimate and the more recent work by Bayer et al. [2007]. These authors' findings suggest homeowners have a marginal willingness to pay that, on average, was approximately a quarter what Black found.[35] While our estimates of the marginal willingness to pay are consistent with the Bayer et al. estimates, as we discuss below, we believe a full analysis would need to take account of the roles of additional locational distinctions for community selections.

Assume each household selects an amount of housing and a numeraire good. Homogeneous housing is available in each of a finite number of communities with different amounts of local public goods. For our application, public education is the primary public good considered by households and the communities correspond to school districts. To describe the logic of the model, suppose the household decision process takes place in two stages. First, a household evaluates what the consumption choices would be if each community was selected, conditional on the housing price and local public goods in each location. This decision includes the selection of an amount of homogeneous housing in each location. Then, in this conceptual description of the process, suppose now a household considers the level of well-being realized in each community with a fixed income and selects the community that provides the highest utility. The numeraire good in this simple description of the choice

[35] More recently, Kuminoff and Pope [2010] have extended the literature further, considering micro-housing sales data and replicating the Black methodology with the Bayer et al. instrument strategy for ten markets. They evaluate the federal No Child Left Behind program and find hedonic measures of the average value for the marginal WTP for improved school quality are four times larger than property value capitalization measures.

process is determined by default. The CES specification for the indirect utility function is given in Equation (13).

$$V_j = \left[\beta q_j^{\ \rho} + \left[e^{\frac{m^{1-\eta}-1}{1-\eta}} e^{\frac{-aP_j^{\varepsilon+1}-1}{1+\varepsilon}} \right]^{\rho} \right]^{1/\rho} \tag{13}$$

The subscript j indexes school districts. The term, q_j corresponds to a measure of public school quality and p_j the price of a standardized unit of housing in community j. η and ε correspond to the income (m) and price (P) elasticities, respectively. We have omitted a subscript for households. As with Sieg et al., the model is estimated using moments based on the percentiles for the income and housing expenditure distributions predicted by the model in comparison to what is observed for each school district given information about prices and the public goods (q_j values).

Seven moment conditions are used in estimation – three income quantiles (25^{th}, 50^{th}, and 75^{th}), three housing quantiles (25^{th}, 50^{th} and 75^{th}), and one expression to describe a measure for the public good index. The ascending bundle condition together with estimates for the housing prices can be used to derive the shares of all households in each community. The expressions for these equilibrium shares can be inverted (numerically) to derive estimates for the public good index that are a function of the model's other parameters. We stack the seven moment conditions and use the ranks of the prices as instruments for the endogenous price and public good measures. We follow Kuminoff [2009] and use Chebyshev polynomials in ranks as instruments.[36]

Data Construction, Education Quality, and Housing Price

The study area is Maricopa County, AZ, USA and includes the Phoenix MSA. Our analysis considers the 46 school districts in this county containing a full set of school quality records and census data. These districts are highlighted in Figure 1 and comprise most of the Phoenix MSA excluding Indian Reservations and uninhabited mountain and desert preserves. At the time of the 2000 Census, the Phoenix MSA had a population of over 3.2 million and was growing rapidly. The population in 2008 was nearly 4.3 million people. Our study considers the period from 2003 to 2006 (largely before the collapse of the housing market in the area). Our data include all of the single family housing transactions for the years 2003

[36] Thanks are due to Nicolai Kuminoff who developed the Matlab Code for the Sieg et al. estimator.

through 2006 as well as school quality data derived from annual school report cards for each school year between 2003 and 2006. Information on income and the population in each school district were developed using the block group SF3 file from the Census 2000 public data.

School quality is measured using the Arizona Department of Education School Report Cards. These reports are published for each school and are mandated by the No Child Left Behind Act. Test scores are available from Arizona's Instrument to Measure Standards (AIMS) test administered to students each spring. These reports contain test scores as well as the number of students in each school, the number of teachers, and teacher aides. The summary of the scores differentiate English versus non-English speaking students and are reported for grades 2 through 12 separately for categories of mathematics, language, and reading for 2003 through 2006. Our analysis considered the test results for grades 2 through 8 and was restricted to the scores for English-speaking students.

We developed measures for the average number of students, teachers, and aides for each grade/district/year combination as well as a measure of the average math, reading, and language score for each grade/school/year combination. The individual school level data were aggregated based on the school's district, the test type, the grade, and the year. Using these averages, we constructed measures for average student/teacher ratio as well as average student/teacher-aide ratio. Considering only the cases with complete records our sample consists of 3711 unique combinations of grade, district, year, and test score.

The measure of school quality used in the estimation of the sorting model is based on predictions from a model of these test scores. It is an index of performance that is associated with observable factors contributing to test scores. We hypothesize that test scores are related to the type of test, the grade level, the year, and the average number of teachers and aides present in each school. The specific measures used in the model are the averages for the student/teacher and student/teacher-aide ratios.

Table 1 provides the results for the model used to construct our school quality index. The estimates are consistent with the literature and our prior hypotheses about the importance of small class sizes. Higher student teacher ratios are associated with lower performance after controlling for the grade, test subject, and year.[37] Increases in the ratio of students to teacher-aides increases test performance. Although this finding may seem counterintuitive, there are a number of potential explanations. One would suggest that this result simply reflects the importance of teachers over aides. That is, the primary factor giving rise to improved test performance is the amount of teacher time devoted to students. If we assume there is a minimum number of staff required for class management

[37] The fixed effects for 2003, grade 2, and language test were omitted in these specifications so the effects that are measured are relative to these reference points for each variable.

and this threshold is met in all cases, then a *lower* number of aides for a given number of teachers would be consistent with an increase in scores because it implies that the class management threshold is more likely to be met with increased amounts of teacher time over that of aides. This outcome would imply increasing the number of aides at the cost of teachers would not enhance school quality. These specific features are important to policies that are associated with teachers' aides. Our example involves reductions in teachers and relies on the effect of these reductions on teaching outcomes and assessments of education quality. Thus, what is important for our purpose is the ability to capture their effects on test performance. While the R^2 score is low the model does capture effects of teachers, recognizes differences for math and reading tests, and thus seems a credible basis for describing how test scores would respond to reductions in teachers at the school level. As a result, we adopt the predicted test score for the language test during the 2003 school year for grade 2 as our index of school quality for each district.

The price index for the homogeneous unit of housing is developed following the framework outlined in Sieg et al. [2002] using a transaction database for residential housing sales in Maricopa County. A log-linear price function is used to estimate district housing price indexes controlling for the attributes of each house sold and the year sold. The price indexes are derived from a set of fixed effect terms, one for each school district. For the period we considered, before the dramatic shift in the housing market, we felt that following Poterba [1992] and annualizing housing prices at a rate of 11% was consistent if not conservative as a description of buyer and seller expectations. The sample of housing sales includes 406,556 transactions between 2003 and 2006. It has all the transaction records for single family homes falling in the 46 school districts. Table 2 reports the estimated price equation. All of the school district fixed effects are significant. The estimated coefficients for the district fixed effects are used to construct price indexes. Their rank, generally, is in agreement with the rank of the districts based on the index of school quality. This consistency is what would be expected based on the ascending bundles condition. It offers confirmation for both the joint decisions in selecting the school districts as a reasonable characterization for the community and for our simple model relating test scores to students per teacher and using it as a gauge of school quality at the primary school level.

The final components of our data are the measures for population in each school district and the income quantities for each school district. These are developed using the block group 100% sample from Census 2000 data reported in the SF3 series of tables. The primary variables are total population counts within each district and the income classified into one of 16 distinct categories with the lowest ranging from $0 to $10,000 and the top category unbounded above $200,000. To construct a spatially consistent measure by school district, the population counts for each block group are weighted based on the area falling within each school district. For example, a block group with half of the area falling in one school district and half the area falling in a second school district would be divided so that the population of people is split between the two districts. By using the total population in each district we construct the share of total population across all districts. This was also the approach used to allocate the count of households in each income bin.

A map of the block groups overlaid on our school districts is shown in Figure 2. Interval censored regression was used to estimate the mean and variance of a log-normal distribution based on the Census income categories. Using the results from the estimated distribution, the estimates for the 25^{th}, 50^{th}, and 75^{th} income quantiles for each school district are recovered. Summary statistics for the school quality, housing price, housing expenditure, household income, and demographic measures for the population are provided in Table 3.

Table 4 presents the generalized method of moments (GMM) estimates for the preference parameters. The estimates for both the price and income elasticities for housing demand are consistent with the literature. The estimate for the rho (ρ) parameter is consistent with satisfying the single crossing property. The school quality measure is also a significant determinant of households' decisions about community location. Thus, the base model offers a plausible description of the link between community choice and local school quality and can serve to illustrate the PE/GE distinction.

It is also possible to develop some intuition about the model by comparing how our estimates for the marginal willingness to pay for a unit change in school quality align with the literature. Our results imply a range of annual values from \$40 to \$83 a year (in 2003 dollars). Bayer et al. [2007] compare hedonic property value estimates and results for a horizontal sorting model applied to the 1990 PUMA data for the six counties in the San Francisco Bay Area (including Alameda, Contra Costa, Maria, San Mateo, San Francisco, and Santa Clara counties). Two sets of estimates are reported in their study. The first of these estimates the effect of a one standard deviation change in average test scores using a hedonic price function using their boundary fixed effects to control for neighborhood effects (i.e. the demographic attributes of the neighbors such as education and race). Their estimates are approximately a quarter the magnitude estimated by Black [1999] and range between \$14 and \$44 per month, with the variation based on the definition of the boundary and whether neighborhood socio-economic characteristics are included in the model. Their preferred estimate was \$17 per month.

To compare this estimate with our results we need to adjust for price changes in residential housing, convert our measure to one in standard deviation units, and compute the monthly equivalent value.[38] After developing these adjustments, the range of estimates for the marginal willingness to pay implied by the parameter estimates of our model is \$5.24 to \$10.86. These estimates are in monthly 1990 dollars for a one standard deviation change in test scores. Given the differences in household incomes between the two areas, they are remarkably close to the Bayer et al. estimates. Based on this criterion, our model provides estimates for both the conventional parameters used to describe housing demand and the relative importance of school quality to households' selections of homes that are consistent with the literature.

VI. A Policy Simulation to Illustrate PE and GE Benefit Measures

For our policy simulation we use a trivariate distribution with income, the taste parameter for schooling, and an assumed number of children per household. We allow for the fact that changes in resources can cause more children to be located in districts with greater resources and yet the ultimate educational outcomes may be inferior. That is, as households with children move to the districts with more resources the students per teacher increase and the school quality declines. To illustrate how recognition of these types of feedbacks can influence outcomes we compare a Nash equilibrium, where households move recognizing what others will do, with the case where households only consider the initial disruption to school quality and the effects it has on prices as households sort. In the Nash equilibrium, both school quality and housing prices are jointly determined in equilibrium. For the second case, only the prices change as a result of the equilibrium sorting. School quality is also a byproduct of movement but only the prices are recognized by households as they move.

To illustrate the effects of market and non-market feedbacks for PE versus GE measures of changes in resources for social programs, we selected a real policy change that involved reductions in state support for local education in Maricopa County. Based on reports distributed through the Arizona Education Association in April 2009, over 1600 teachers were fired in the county. These effects were unevenly distributed throughout the county's school districts.

Two simulations were developed. The first considers market and non-market feedback effects. It introduces the reductions in the teachers for each district. Using our estimated school quality function given in Table 1 we estimate, with existing students, the reduction in test scores. Recall increased student/teacher ratios reduce test scores. With no moving costs any reduction in school quality creates incentives for some households to move. These changes, in

[38] The housing component of the CPI was 128.5 in 1990 and 184.0 in 2003. As Table 3 suggests, the standard in test scores for our sample was 2.25. This scaling the endpoints of our range by 2.25 × (128.5/184.0) × (1/12) will provide comparable measures.

turn, alter housing prices and create more incentives for households to resort until housing price adjustments imply there are no further gains to movement.

Three aspects of the development of our policy scenarios should be noted before turning to our results. The first concerns replicating the benchmark equilibrium. Our index of educational quality relies on the estimates for test scores as reported in Table 1. To assure they are consistent with the benchmark equilibrium and with our estimates for equilibrium housing prices, we begin our analysis by simulating the benchmark case and adjusting the intercept of the education function so we exactly match the population shares in each school district. The second issue concerns the assumed family size. Here we rely on Census estimates for the mean and variance of family sizes.[39] Finally, our analysis generated one million values for income, the taste parameter for education, and family size using a trivariate normal distribution based on the estimates for these parameters. The initial benchmark solution used to establish the corrections to the intercepts for the school quality equation assigns each household to a school district.

The first simulation considers the Nash equilibrium where households recognize the effects of sorting on both price *and* school quality through the children assigned to each simulated household. When households move the student/teacher ratio changes and school quality adjusts accordingly. In our analysis, the only exogenous change is the reduction in the number of teachers in each district based on the policies for Maricopa County with proposed cuts in teachers by district. Table 5 presents a selection of results for a few school districts for this simulation. The full details for all districts are in Appendix A (see Table 1A for the first scenario with non-market feedbacks). The first column in Table 5 provides the school district name. The second column is the percentage reduction in teachers as part of the budget reduction. Out of 46 districts, 12 lost teachers with cuts ranging from 3% to 27% of the teaching staff. Columns 3 and 4 report the proportionate change in the equilibrium price and the index of public goods. We selected seven districts with varying effects of the policy to illustrate the results. The results for willingness to pay are grouped into GE and PE with the GE marginal willingness to pay (MWTP) corresponding to the marginal willingness to pay (i.e. V_q/V_m) for school quality evaluated at the new GE prices and quality. The PE corresponds to the same concept evaluated at the initial values. These are reported as a gauge of the effects of the point of evaluation for these marginal values.

Our analysis follows the Smith et al. [2004] convention and treats households as renters, so the capitalization effects due to price changes from the

[39] The mean for family size was 3.116. We subtracted 2 to reflect parents and restricted the children to be a positive value or zero. The variance was 0.0326. Family size was assumed to be negatively correlated with income and independent of taste for education.

initial housing assignment accrue to absentee land owners. Consider first the proportionate change in school quality and housing prices created by comparing the benchmark solution to the new equilibrium. Several important results emerge from this comparison. First, averaging across households in a school district, it appears that everyone loses from cuts in teachers even though the cuts are restricted to approximately a quarter of the districts. Because there are no costs to moving, the model assumes households will adapt to attempt to find districts with better test scores after the cuts. However, these actions result in a spreading of the "pain" through price increases in those districts where school quality increases slightly. When quality declines housing prices may decline.

We find that equilibrium schooling quality declines in most places. The quality changes because of two considerations. The policy reduces teachers and household sorting leads to a non-market feedback. Students move into districts with more teachers, raising the student teacher ratio and reducing test scores. This effect also serves to spread the losses from the selective reduction in teachers throughout the county. For the seven districts we profiled, there is little difference between the marginal willingness to pay measures based on the point of evaluation. The primary impact of the GE and PE perspective can be seen through the difference in the WTP measures from a GE verses a PE perspective (averaged across households based on their original school district in the baseline solution of the model before the policy is implemented).

The PE results consider only the change in school quality. They include districts that would appear to have small annual gains and others with large losses, amounting to over $160 per month. These gains can be traced to situations where school quality increased slightly and the housing price increases are ignored. Palo Verde and Roosevelt experience small gains from the PE perspective because we allow for sorting after the change in teachers but do not change housing prices in the computation of the willingness to pay. In all other cases in Table 5, the school quality declines and the distinction between PE and GE arises in whether the housing price change is recognized. As the details of the computations suggest, we could also define other benefit measures, assuming, for example, no sorting and using the school quality based on the initial number of students with whatever number of teachers was implied by the policy change. This possible alternative highlights how the distinction between the definition of the policy to be evaluated and how it is represented in the model can matter for the interpretation of the results from PE and GE benefit measures.

It would be difficult without the structural model to anticipate some of these results. For example, Paradise Valley (a very expensive area north of Phoenix) experiences a cut in teachers, reduction in school quality but an *increase* in housing prices so the GE loss is larger than the PE, a contrast with most of the other findings where housing price reductions serve to reduce the net impact of the decline in school quality. These types of responses are most evident in the Higley Unified and Cave Creek districts shown in Table 5. Overall, the GE losses range from \$41 to \$91 per month.

Table 6 selects the same school districts (results for all are in Table 2A in Appendix A) and compares the GE and PE estimates for a solution based on sorting on price alone with school quality exogenous. These are given in the four columns after restating the school district name and percentage cut in teachers. In this case, teacher cuts result in school quality declines using the initial distribution of students and reduced number of teachers as the assumed base for welfare computation. Households resort and this sorting process affects housing prices but school quality is assumed to be determined by the initial distribution of students.

In columns 7 through 10 of Table 6, we round the estimates of WTP to whole dollars from the Nash solutions (in Table 5) and include them for comparison. The Nash adjustment in school quality increases the estimates for GE losses (in absolute magnitude). The PE losses are smaller (in absolute magnitude) for the Nash than the solution that does not consider endogeneity in the non-market outcomes caused by sorting. GE losses are smaller, as might be expected when households are assumed to anticipate how their own and others' behavior will affect school quality. Consider, for example, Palo Verde school district or Roosevelt Elementary. The PE measures derived from the Nash equilibrium are positive, whereas they are negative when adjustment is based on price. In these cases, the assumed Nash adjustment leads to an increase in school quality, whereas sorting based on price alone implies a very small decline in school quality in one case and no change in another.

Overall, our sorting example illustrates three features of the comparison of GE and PE measures of the willingness to pay to avoid declines in school quality. First, it is possible to exploit revealed preference logic to develop models capable of reflecting multi-market adjustment in response to policies affecting social programs. Second, when the programs exist in different jurisdictions both market and non-market adjustments are possible. The non-market feedbacks are likely to be *more important* to discrepancies between PE and GE measures than price effects alone. Finally, measures of marginal willingness to pay were not as sensitive to the point of evaluation as comparisons of PE versus GE willingness to pay might lead an analyst to speculate would be the case.

VII. Summary and Research Ahead

This paper has summarized definitions for PE and GE welfare measures when policy is assumed to affect only market goods. We generalized these definitions to consider market and non-market goods and outlined two modeling strategies for measuring the importance of GE effects. Finally, we illustrated how one of these frameworks, a locational equilibrium model, could be used to estimate PE and GE welfare measures for local public education policies as an example of a social policy. While our estimates closely match the literature relevant to the application, they are intended here simply as an example. A more complete analysis would need to consider air quality, crime, and a variety of other spatially delineated factors that influence neighborhood choices.

Several research issues are "buried" in the details of model implementation that should be considered in future research. We highlight three here: measures for the "amounts" of social programs; revealed preference and the nonuse values for changes in social programs; and the extent of the market for social programs. We close with a short discussion of each issue.

A. Quantity Measures for Social Program Outputs

Our example of education policies as a social program focused on one measure of the output – school quality measured by test scores. If the objective of public education is to assure an informed electorate so that a democracy provides "better" decisions, then the relationship between test scores at the primary school level and an "informed electorate" is certainly not clear. If we believe education helps to avoid other social problems or enhances the likelihood for good social outcomes on a number of dimensions (i.e. crime, teenage pregnancy, childhood poverty, etc.) as Heckman has argued, then it seems reasonable to assume these effects are unlikely to be captured by the gains realized by individual households who seek to enhance the private skills of their children. How these individual choices "add up" to transform the collective outcome is certainly omitted from the types of models we discuss here. Of course, to deal with them we need a framework that links observable outcomes to the sequence of steps the improvements in the quality of local education will facilitate.

The task becomes more complex for social programs with limited direct private benefits. These issues must be addressed to quantitatively "scale" the output in a CGE setting. Measures of nonuse value or the benefits from purely public outcomes of social programs cannot be recovered from revealed preference models. We need to consider integrating results from stated preference analyses with revealed preference measures of the tradeoffs people would make for the private benefits associated with each specific program.[40]

[40] See Carbone and Smith [2010] for further discussion of the first point.

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

B. Nonuse Values for Social Programs

Environmental economists have been concerned about people who care about environmental resources that they may never want to "use". These preferences need not stem from an altruistic motive directed at the current or some future generation. It is certainly possible to consider preferences for a society that sustains social programs. It may be the case that individuals would make decisions (if they were available) to give up resources for these outcomes. We simply do not observe them. Once again this raises issues about how we measure tradeoffs to calibrate preferences. In most CGE models for market goods, we assume these marginal tradeoffs are revealed through ideal markets. In these cases we simply do not observe them.

C. Extent of the Market

At an aggregate level, judging the importance of GE effects will depend on these tradeoff measures *and* the extent of the market. That is, how many people have them? Such questions do not come up for market transactions because expenditure flows allow the analyst to scale-up consumption levels and create the representative consumer. A comparable process can be used for user values for non-market goods on the revealed preference logic. Neither is available for social programs that largely resemble the concepts classified as nonuse services. Both the characterization of the tradeoffs with market goods and "the extent of the market" (or the aggregate resources that would be made available by people who would make these tradeoffs) determine the importance of the GE effects of policies influencing these non-market goods.

These issues can be addressed. Some progress has been made for policies that are intended to change environmental resources. However, the record is much more limited with social programs. A prudent starting point would be to begin with social policies that can be addressed with some variant of the revealed preference logic, such as illustrated here with our sorting model for education and progressively extend the analysis to other situations with large benefits hypothesized to be associated with public good effects as we gain experience with the narrower set of problems where the revealed preference logic can be exploited.

Table 1. School Quality Regression Model

Variable	Estimate	Std Err	t-stat
Student/Teacher	-0.2493	0.1074	-2.32
Student/Teacher Aide	0.0598	0.0061	9.81
Grade 3	-0.6073	0.8297	-0.73
Grade 4	0.3291	0.8293	0.40
Grade 5	0.1127	0.8285	0.14
Grade 6	1.1122	0.8293	1.34
Grade 7	2.7064	0.8317	3.25
Grade 8	2.1403	0.8333	2.57
Math	5.0642	0.5442	9.31
Reading	1.2624	0.5444	2.32
Year 2004	-0.3870	0.6327	-0.61
Year 2005	-1.5289	0.6308	-2.42
Year 2006	-2.6884	0.6290	-4.27
Constant	48.4230	2.0401	23.74

R-square= .0594
N=3711

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

Table 2. Fixed Effect Hedonic Property Model for Maricopa County School Districts 2003-2006

Variable	Estimate	Std Err	t-stat	Variable	Estimate	Std Err	t-stat
Lot Acres	0.3020	0.0030	100.66	District 18	8.9333	0.0053	1676.05
Square Feet (100s)	0.0569	0.0003	182.24	District 19	8.7462	0.0078	1114.49
Stories	-0.1452	0.0013	-109.92	District 20	9.1150	0.0055	1665.36
Bathrooms	0.0670	0.0012	57.27	District 21	8.7727	0.0055	1585.03
Age	-0.0083	0.0001	-71.06	District 22	8.8561	0.0065	1372.61
Lot Acres Sq	-0.0234	0.0006	-39.19	District 23	8.8902	0.0053	1677.99
Square Feet (100s) Sq	-0.0005	0.0000	-90.50	District 24	8.8079	0.0056	1585.03
Age Sq	0.0001	0.0000	31.56	District 25	9.3861	0.0074	1267.62
Garage	0.0411	0.0020	20.41	District 26	8.9157	0.0049	1804.82
Pool	0.0758	0.0011	67.55	District 27	9.6041	0.0604	159.01
Year 2004	0.1116	0.0019	59.03	District 28	8.2523	0.0282	292.52
Year 2005	0.3936	0.0019	209.71	District 29	8.5662	0.0136	630.22
Year 2006	0.5521	0.0020	281.73	District 30	8.3126	0.0116	717.23
District 1	8.2976	0.0353	235.27	District 31	9.1222	0.0086	1066.71
District 2	8.8099	0.0067	1322.38	District 32	8.2997	0.0226	367.28
District 3	7.9440	0.0279	285.09	District 33	9.1620	0.0051	1789.31
District 4	8.7445	0.0057	1534.37	District 34	8.8030	0.0056	1574.50
District 5	9.0810	0.0114	798.06	District 35	8.9441	0.0050	1801.12
District 6	8.6239	0.0055	1556.11	District 36	9.0546	0.0076	1196.21
District 7	8.7563	0.0057	1545.80	District 37	8.7389	0.0065	1343.83
District 8	9.2811	0.0058	1606.96	District 38	8.7416	0.0098	889.23
District 9	8.9473	0.0049	1826.41	District 39	8.7555	0.0052	1683.18
District 10	9.1132	0.0072	1258.24	District 40	8.9351	0.0121	737.79
District 11	8.9726	0.0049	1833.24	District 41	9.4296	0.0054	1741.63
District 12	8.8057	0.0047	1866.59	District 42	9.0172	0.0061	1467.58
District 13	9.3064	0.0074	1264.40	District 43	8.7829	0.0079	1112.95
District 14	8.7443	0.0065	1339.52	District 44	8.9009	0.0055	1633.04
District 15	7.9725	0.0290	274.85	District 45	8.6870	0.0110	791.37
District 16	8.9419	0.0050	1781.48	District 46	8.3221	0.0305	273.00
District 17	8.7980	0.0058	1504.79				

R-square=.999

N=406,556

Table 3. Summary Statistics for Characteristics of Maricopa County School Districts Used in Sorting Model

Variable	Mean	Std Dev	Min	Max
Price Index Rank	23.5	13.4	1.0	46.0
Population Share	0.0217	0.0290	0.0003	0.1437
Test Score	48.43	2.25	43.21	56.26
Price Index	8.83	0.35	7.94	9.60
Income 25th Pct	23,307	8,133	11,526	44,099
Income 50th Pct	39,307	12,882	21,056	77,497
Income 75th Pct	66,501	21,190	36,960	136,189
House Price 25th Pct	18,772	9,221	4,950	58,300
House Price 50th Pct	25,169	11,019	7,810	59,400
House Price 75th Pct	34,428	16,003	9,900	85,250
Household Size	3.12	0.89	2.06	6.90
Students	12,562	13,972	21	59,701
Teachers	678	749	3	3,051
Teacher Aides	212	363	1	2,174
# Schools	14	16	1	77

Number of districts=46

Table 4: GMM Estimation Results for Household Preferences*

Variable	Estimate	Std Error	t-stat
std dev for Ln(inc)	0.4332	0.0029	148.8200
mean for Ln(inc)	10.5296	0.3154	
mean for taste par.	0.9092	0.0764	11.9080
std dev for taste par	0.1806	0.0079	22.8610
lambda	-0.2758	0.0025	-112.4700
income elasticity	0.9214	0.0015	630.6300
price elasticity	-0.4781	0.0370	-12.9280
beta	1.3215	0.0254	52.0960
rho	-0.0438	0.0005	-88.6630
q_initial	46.4400	0.1412	328.8600

*** Standard errors are generated from bootstraps using 5 iterations, std deviation for mean income based on census data N=1,000,000for simulation**

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

Table 5: **General and Partial Equilibrium Measures for Willingness to Pay: Nash Equilibrium** [a]

School District	Loss of Teachers	Change in Housing Price	Change in Public Good Index	GE WTP	GE MWTP	PE WTP	PE MWTP
Palo Verde Elementary District	0	0.00687	0.0002	-490.16	40.59	6.30	40.64
Queen Creek District	27	0.00062	-0.0109	-501.36	41.34	-455.45	40.8
Litchfield Elementary District	18	-0.00242	-0.01342	-518.47	50.90	-786.98	50.07
Higley Unified District	21	-0.00234	-0.01371	-612.95	55.39	-1108.71	54.22
Roosevelt Elementary District	0	0.00695	0.0001	-514.48	45.29	4.25	45.35
Paradise Valley District	9	0.00078	-0.00933	-698.2	63.32	-553.28	62.81
Cave Creek District	18	-0.00880	-0.01914	-719.21	65.39	-1344.35	64.03

[a] See Appendix Table 1A for full details.

Table 6: **Comparison of GE and PE- Nash Equilibrium Versus Price Only Solution**

School District	Loss of Teachers	Price Only: Change in Price	Price Only: Change in Public Good	Price Only: GE WTP	Price Only: PE WTP	Nash: Change in Price	Nash: Change in Public Good	Nash: GE WTP	Nash: PE WTP
Palo Verde Elementary District	0	0.00681	-	-494	-	-0.00687	0.00020	-490	6
Queen Creek District	27	0.00085	-0.01059	-504	-441	-0.00062	-0.01090	-501	-455
Litchfield Elementary District	18	-0.00252	-0.01364	-523	-766	-0.00242	-0.01342	-518	-787
Higley Unified District	21	-0.00146	-0.01269	-618	-1132	-0.00234	-0.01371	-613	-1109
Roosevelt Elementary District	0	0.00693	-	-517	-	0.00695	0.00010	-514	4
Paradise Valley District	9	0.00030	-0.00995	-705	-561	0.00078	-0.00933	-698	-553
Cave Creek District	18	-0.00860	-0.01904	-726	-1338	-0.00880	-0.01914	-719	-1344

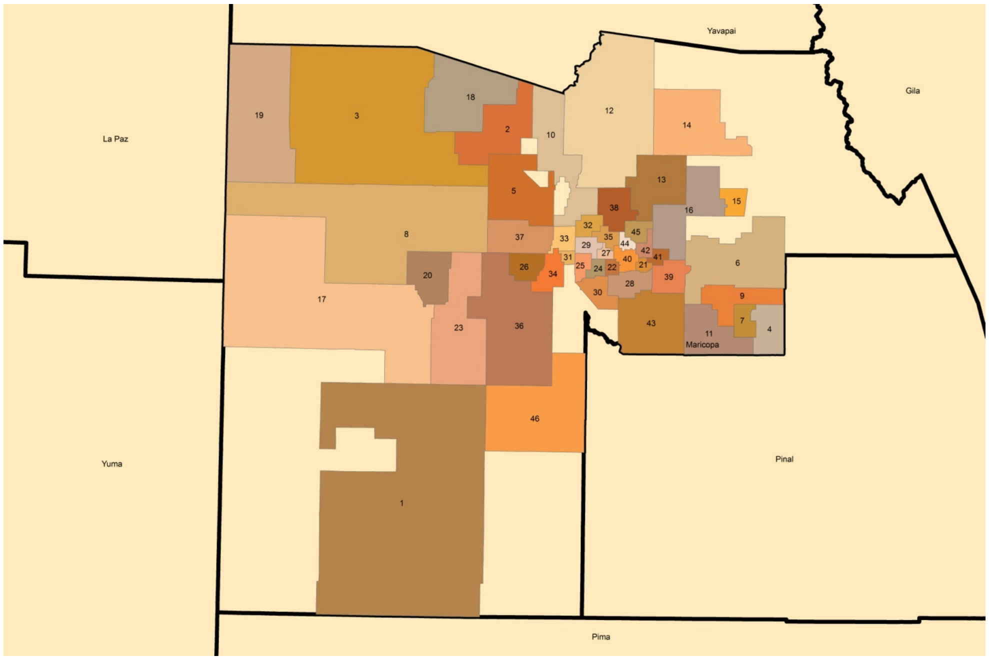

Figure 1: School Districts in Maricopa County, AZ, USA

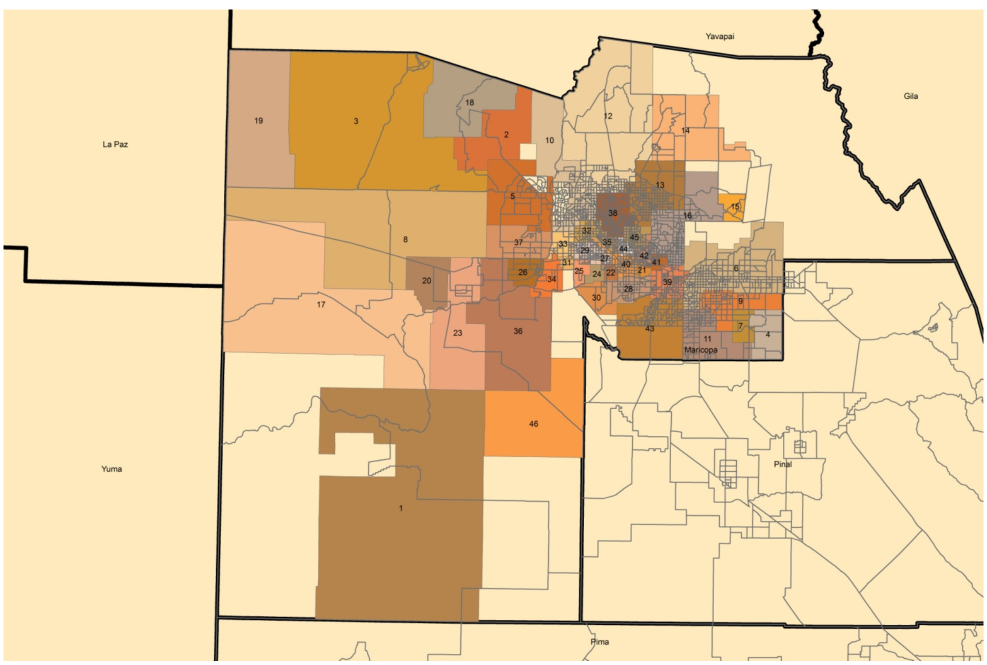

Figure 2: Census Block Groups Overlaid with Maricopa School Districts

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

Appendix A

Table 1A: General and Partial Equilibrium Measures of Willingness to Pay for All School Districts: Nash Equilibrium

schoold	district	loss	gewtp	pewtp	mwtpold	mwtpnew	dq	dp
1	ARLINGTON ELEMENTARY DISTRICT	0	-490.38943	-13.527759	40.259444	40.230033	-.0004018	.0067556
2	GILA BEND DIST.	0	-504.76866	-7.306073	44.433164	44.390347	-.0001492	.006866
3	MORRISTOWN ELEM. DIST.	0	-491.40097	-22.798008	40.803003	40.780765	-.0002999	.0066339
4	AGUILA ELEM. DIST.	0	-510.08132	-69.735639	43.456627	43.493236	-.0015271	.0059797
5	PALO VERDE ELEMENTARY DISTRICT	0	-490.16196	6.3026169	40.641551	40.585726	.0001972	.0068686
6	NADABURG DIST.	0	-490.89503	-21.185502	41.68903	41.666205	-.000534	.0064792
7	WILSON ELEMENTARY DISTRICT	0	-502.50206	-5.4290461	43.831858	43.787035	-.0001104	.0067001
8	MURPHY ELEMENTARY DISTRICT	0	-500.87852	-102.44376	42.3859	42.458013	-.0060759	.0033984
9	BUCKEYE ELEMENTARY DISTRICT	0	-498.98996	-410.17369	41.890434	42.31981	-.0119292	.0000544
10	WICKENBURG DIST.	0	-508.8757	-538.0534	42.644868	43.219903	-.0130557	-.0005925
11	QUEEN CREEK DIST.	27	-501.36401	-455.44873	40.875015	41.341651	-.0109218	.000619
12	RIVERSIDE ELEMENTARY DISTRICT	0	-495.7499	-265.46507	39.959439	40.208151	-.0060156	.0034135
13	FOWLER ELEMENTARY DISTRICT	0	-515.35501	-281.65795	42.6209	42.89145	-.0003772	.0066804
14	AVONDALE ELEMENTARY DISTRICT	5	-517.90611	-104.60041	43.281137	43.347715	.0002133	.0070217
15	ISAAC ELEMENTARY DISTRICT	0	-517.63118	-.5061104	44.289605	44.235534	-.0000109	.00689
16	ROOSEVELT ELEMENTARY DISTRICT	0	-514.47881	4.2486439	45.352614	45.291817	.0001032	.0069537
17	CARTWRIGHT ELEMENTARY DISTRICT	0	-511.40305	-.95035779	46.591663	46.535288	-.0000201	.0068764
18	LAVEEN ELEMENTARY DISTRICT	0	-507.40299	-78.414726	46.8302	46.86207	-.0015017	.0059092
19	TOLLESON ELEMENTARY DISTRICT	0	-512.15242	-158.53226	47.488125	47.611333	-.0018939	.0056541
20	GLENDALE ELEMENTARY DISTRICT	0	-511.37385	-58.746621	47.645785	47.653667	-.0026111	.0051872
21	PENDERGAST ELEMENTARY DISTRICT	0	-512.96333	-48.290093	48.692144	48.687146	-.0009761	.0062443
22	DYSART DIST.	0	-514.61709	-68.425549	49.535268	49.552494	-.0013299	.0060052
23	LITTLETON ELEMENTARY DISTRICT	0	-513.22004	-85.572521	49.769886	49.806584	-.001277	.0060482
24	ALHAMBRA ELEMENTARY DISTRICT	0	-512.75566	-103.13234	49.794211	49.850256	-.0008173	.0063849
25	LIBERTY ELEMENTARY DISTRICT	0	-512.04251	-491.54559	49.42409	49.918331	-.0082804	.0011682
26	LITCHFIELD ELEMENTARY DISTRICT	18	-518.4737	-786.97677	50.071393	50.900916	-.0134275	-.0024214
27	WASHINGTON ELEMENTARY DISTRICT	0	-534.86717	-841.59576	50.607628	51.492398	-.0145812	-.0031869
28	MESA UNIFIED DIST.	7	-582.19313	-975.10337	52.884543	53.910709	-.0165885	-.0046419
29	HIGLEY UNIFIED DISTRICT	21	-612.95549	-1108.7067	54.217491	55.386593	-.0137122	-.0023404
30	SADDLE MOUNTAIN UNIFIED DISTRICT	0	-607.46621	-1112.5434	53.263926	54.428495	-.0115285	-.0006144
31	GILBERT DIST.	18	-622.49038	-802.51034	54.648602	55.468738	-.011095	-.0002633
32	PEORIA DIST.	15	-631.55101	-356.86198	55.824382	56.142021	-.0060616	.0037086
33	CHANDLER DIST.	0	-636.70084	-163.521	56.888237	56.984643	-.0021709	.0068742
34	DEER VALLEY DIST.	4	-643.29688	-248.6952	58.135862	58.326143	-.0046049	.0048577
35	TEMPE ELEMENTARY DISTRICT	0	-643.33356	-206.41051	58.357359	58.496803	-.0092786	.0008754
36	PHOENIX ELEMENTARY DISTRICT	0	-642.99049	-308.76799	58.668589	58.921068	-.0128077	-.0021962
37	BALSZ ELEMENTARY DISTRICT	0	-649.43834	-503.58203	59.452283	59.922008	-.0083008	.0015706
38	CREIGHTON ELEMENTARY DISTRICT	3	-654.92593	-773.52231	59.562762	60.329983	-.0119121	-.001586
39	KYRENE ELEMENTARY DISTRICT	0	-670.03012	-742.27325	60.604916	61.333758	-.0074373	.0024637
40	OSBORN ELEMENTARY DISTRICT	0	-680.54862	-550.46333	61.456541	61.968451	-.0079016	.0020439
41	PARADISE VALLEY DIST.	9	-698.19829	-553.27627	62.808903	63.316734	-.0093285	.0007782
42	CAVE CREEK DIST.	18	-719.20999	-1344.3497	64.026544	65.392575	-.019143	-.0087977
43	FOUNTAIN HILLS DIST.	19	-721.75168	-1519.4717	63.066454	64.609226	-.0210388	-.0106322
44	MADISON ELEMENTARY DISTRICT	0	-749.74946	-1909.1	63.810503	65.762463	-.0256822	-.0150985
45	SCOTTSDALE DIST.	13	-919.3232	-1925.9011	69.035853	70.970869	-.0071384	.0059341
46	MOBILE ELEM. DIST.	0	-1093.2008	6.6730281	83.561564	83.312242	.0000683	.0155473

Table 2A: General and Partial Equilibrium Measures of Willingness to Pay: Equilibrium for Teacher Cuts Based on Prices Only

schoold	district	loss	gewtp	pewtp	dq	dp
1	ARLINGTON ELEMENTARY DISTRICT	0	-494.98017	0	3.06e-16	.0070134
2	GILA BEND DIST.	0	-508.76134	-1.7742719	-1.52e-16	.0069965
3	MORRISTOWN ELEM. DIST.	0	-495.02173	-11.855071	1.42e-16	.0068362
4	AGUILA ELEM. DIST.	0	-513.44588	-12.164209	0	.0068111
5	PALO VERDE ELEMENTARY DISTRICT	0	-493.35632	-.23503178	0	.00681
6	NADABURG DIST.	0	-493.89379	-1.0731877	-1.40e-16	.0068029
7	WILSON ELEMENTARY DISTRICT	0	-505.34552	-.86583073	-4.19e-16	.0067977
8	MURPHY ELEMENTARY DISTRICT	0	-503.76719	-107.64579	-.0061744	.0033826
9	BUCKEYE ELEMENTARY DISTRICT	0	-501.9955	-414.89058	-.0120764	.0000119
10	WICKENBURG DIST.	0	-511.91192	-535.3155	-.0129999	-.0005208
11	QUEEN CREEK DIST.	27	-504.22877	-441.4307	-.0105893	.0008467
12	RIVERSIDE ELEMENTARY DISTRICT	0	-498.50061	-255.71386	-.0057885	.0035806
13	FOWLER ELEMENTARY DISTRICT	0	-518.13883	-271.86577	-.0005601	.0066102
14	AVONDALE ELEMENTARY DISTRICT	5	-520.623	-107.1448	-.0000282	.0069177
15	ISAAC ELEMENTARY DISTRICT	0	-520.3991	-.07489656	1.53e-15	.0069332
16	ROOSEVELT ELEMENTARY DISTRICT	0	-517.31574	-.41510836	-1.27e-16	.0069282
17	CARTWRIGHT ELEMENTARY DISTRICT	0	-514.27833	-.03808399	3.81e-15	.0069278
18	LAVEEN ELEMENTARY DISTRICT	0	-510.1121	-28.644659	-.000335	.0067026
19	TOLLESON ELEMENTARY DISTRICT	0	-514.72394	-102.32895	-.0017937	.0057524
20	GLENDALE ELEMENTARY DISTRICT	0	-514.18766	-90.189852	-.0026468	.0051965
21	PENDERGAST ELEMENTARY DISTRICT	0	-516.64294	-97.890529	-.0019764	.0056231
22	DYSART DIST.	0	-518.97509	-117.57657	-.0013022	.006086
23	LITTLETON ELEMENTARY DISTRICT	0	-517.82753	-84.172384	-.0013037	.0060884
24	ALHAMBRA ELEMENTARY DISTRICT	0	-517.32671	-98.864273	-.0011445	.0062108
25	LIBERTY ELEMENTARY DISTRICT	0	-516.49586	-471.28208	-.0082161	.0012754
26	LITCHFIELD ELEMENTARY DISTRICT	18	-522.89931	-766.35249	-.0136382	-.002515
27	WASHINGTON ELEMENTARY DISTRICT	0	-539.33404	-846.08607	-.0150543	-.0034856
28	MESA UNIFIED DIST.	7	-587.15817	-988.04729	-.0169885	-.0048837
29	HIGLEY UNIFIED DISTRICT	21	-618.3816	-1132.4287	-.0126912	-.0014594
30	SADDLE MOUNTAIN UNIFIED DISTRICT	0	-612.85236	-1136.0491	-.012318	-.0011626
31	GILBERT DIST.	18	-628.01804	-810.71504	-.0100883	.0006059
32	PEORIA DIST.	15	-637.38064	-376.51179	-.0063927	.0035208
33	CHANDLER DIST.	0	-642.98338	-185.23855	-.0025399	.0066567
34	DEER VALLEY DIST.	4	-649.94682	-256.48348	-.0051659	.0044801
35	TEMPE ELEMENTARY DISTRICT	0	-649.84502	-198.31366	-.0092975	.0009487
36	PHOENIX ELEMENTARY DISTRICT	0	-649.34187	-302.56993	-.0125817	-.0019153
37	BALSZ ELEMENTARY DISTRICT	0	-655.82645	-507.61753	-.0081807	.0017603
38	CREIGHTON ELEMENTARY DISTRICT	3	-661.34439	-777.61402	-.0119741	-.0015548
39	KYRENE ELEMENTARY DISTRICT	0	-676.53374	-748.35132	-.0076589	.0023512
40	OSBORN ELEMENTARY DISTRICT	0	-687.2344	-567.67416	-.0082762	.0017916
41	PARADISE VALLEY DIST.	9	-705.34113	-560.92684	-.0099518	.0002959
42	CAVE CREEK DIST.	18	-726.23899	-1337.8835	-.0190385	-.0086028
43	FOUNTAIN HILLS DIST.	19	-728.64363	-1513.999	-.0209653	-.0104676
44	MADISON ELEMENTARY DISTRICT	0	-756.73894	-1916.1204	-.0257835	-.0151063
45	SCOTTSDALE DIST.	13	-923.51859	-1878.0036	-.0047263	.0086774
46	MOBILE ELEM. DIST.	0	-1055.411	0	0	.0149221

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

References

Banzhaf, H. Spencer and V. Kerry Smith. 2007. "Meta-analysis in Model Implementation: Choice Sets and the Valuation of Air Quality Improvements." *Journal of Applied Econometrics*, 22(6): 1013–1031.

Bayer, Patrick, Fernando Ferreira and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy*, Vol. 115(4) August: 588–638.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, Vol. 114(2) May: 577–599.

Bockstael, Nancy E. and Kenneth E. McConnell. 2007. *Environmental and Resource Valuation with Revealed Preferences* (Dordrecht, the Netherlands: Springer.)

Böhringer, Christopher and Thomas F. Rutherford. 2008. "Combining Bottom-up and Top-down" *Energy Economics*, Vol. 30(2): 574–596.

Boyle, Kevin J., Nicolai V. Kuminoff, Christopher F. Parameter and Jaren C. Pope. 2010. "The Benefit Transfer Challenges." *Annual Review of Resource Economics*, Vol. 2 (Palo Alto, CA: The Annual Reviews) pp. 161–182.

Carbone, Jared C. 2005. "Calibrating Non-separable General Equilibrium Models." *Unpublished Working paper*, CEnREP: North Carolina State University, March.

Carbone, Jared C. 2007. "The Properties of Non-Price Equilibria." *Unpublished Working Paper*, Department of Economics, Williams College, August 27.

Carbone, Jared C. and V. Kerry Smith. 2008. "Evaluating Policy Interventions With General Equilibrium Externalities." *Journal of Public Economics*, Vol. 92(5–6) June: 1254–1274.

Carbone, Jared C. and V. Kerry Smith. 2010. "Valuing Ecosystem Services in General Equilibrium." March, *NBER Working Paper* #15844.

Carbone, Jared C., Helm, Carsten and Rutherford, Thomas F. "The case for international emission trade in the absence of cooperative climate policy" *Journal of environmental Economics and management*, Vol. 58(3): 266–280.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Models." *Annual Review of Economics*, Vol. 1 (Palo Alto, CA: Annual Reviews): 451–488.

De Mooij, Ruud A. 2000. *Environmental Taxation and the Double Dividend*, (Amsterdam: North Holland).

Diamond, Peter A. and Daniel L. McFadden. 1974. "Some Uses of the Expenditure Function in Public Finance." *Journal of Public Economics*, Vol. 3(1) February: 3–21.

Epple, Dennis and Glenn J. Platt. 1998. "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes." *Journal of Urban Economics*, Vol. 43(1) January: 23–51.

Epple, Dennis and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdiction." *Journal of Political Economy*, Vol. 107(4) August: 645–681.

Epple, Dennis, T. Romer and Holger Sieg. 2001. "Interjurisdictional Sorting and Majority Rule: An Empirical Analysis." *Econometrica*, Vol. 69(6) November: 1437–1465.

Epple, Dennis, Brett Gordon and Holger Sieg. 2009. "Drs. Mirth and Mills meet Dr. Tiebout: Integrating Location Specific Amenities into Multi-community Equilibrium Models." Unpublished paper prepared for 50th Anniversary Conference of *Journal of Regional Science*, March.

Espinosa, J. Andrés and V. Kerry Smith. 1995. "Measuring the Environmental Consequences of Trade Policy: A Nonmarket CGE Analysis." *American Journal of Agricultural Economics*, Vol. 77(3) August: 772–777.

Farrow, Scott. 1999. "The Duality of Taxes and Tradable Permits: A Survey with Applications in Central and Eastern Europe." *Environmental and Development Economics*, Vol. 4(4): 519–535.

Fernandez, Raquel and Richard Rogerson. 1998. "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform." *The American Economic Review*, Vol. 88(4) September: 813–833.

Ferreyra, Maria Marta. 2007. "Estimating the Effects of Private School Vouchers in Multidistrict Economies." *American Economic Review*, Vol. 97(3) June: 789–817.

Ferreyra, Maria Marta. 2009. "An Empirical Framework for Large Scale Policy Analysis with an Application to School Finance Reform in Michigan." *American Economic Journal, Economic Policy*, Vol. 1(1) February: 147–180.

Freeman, A. Myrick III. 2003. *The Measurement of Environmental and Resource Values: Theory and Methods*, second edition. Washington DC: Resources for the Future.

Fullerton, Don, Yolanda K. Henderson and John B. Shoven. 1984. "A Comparison of Equilibrium Models of Taxation." In *Applied General Equilibrium Analysis* edited by H.E. Scarf and J.B. Shoven. Cambridge University Press.

Fullerton, Don and Gilbert Metcalf. 1997. "Environmental Controls, Scarcity Rents and Pre-existing Distortions." *NBER Working Paper* # 6091.

Goulder, Lawrence H. and Roberton C. Williams III. 2003. "The Substantial Bias from Ignoring General Equilibrium Effects in Estimating Excess Burden, and a Practical Solution." *Journal of Political Economy*, Vol. 111(4) August: 898–927.

Goulder, Lawrence H., Ian W.H. Parry, Roberton C. Williams III and Dallas Burtraw.1999. "The Cost-Effectiveness of Alternative Instruments for Environmental Protection in a Second-Best Setting," *Journal of Public Economics*, Vol. 72(3) June: 329–360.

Hanemann, W. Michael. 1991. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review*, Vol. 81(3 June):635–647.

Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review*, Vol. 54(1): 58–76.

Harrison, Glenn W. and H.D. Vinod. 1992. "The Sensitivity of Analysis of Applied General Equilibrium Models: Completely Randomized Factorial Designs." *Review of Economics and Statistics*, Vol. 74(2) May: 357–362.

Hazilla, Michael and Raymond J. Kopp. 1990. "The Social Cost of Environmental Quality Regulation: A General Equilibrium Analysis." *Journal of Political Economy*, Vol. 98(4) August: 853–873.

Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science*, Vol. 312(5782) June: 1900–1902.

Heckman, James J. 2008. "Schools, Skills and Synapses." *Economic Inquiry*, Vol. 46(3) July: 289–324.

Hoehn, John and Alan Randall. 1989. "Too Many Proposals Pass the Benefit Cost Test." *American Economic Review*, Vol. 79(3) June: 544–551.

Jones, Ronald W. 1965. "The Structure of Simple General Equilibrium Models." *Journal of Political Economy*, Vol. 73(6) December: 557–572.

Just, Richard E., Darrell L. Hueth and Andrew Schmitz. 2004. *The Welfare Economics of Public Policy: A Practical Approach to Project and Policy Evaluation*. Cheltenham, UK: Edward Elgar.

Klaiber, H. Allen and Daniel J. Phaneuf. 2010. "Valuing Open Space in a Residential Sorting Model of the Twin Cities." *Journal of Environmental Economics and Management*, 60(2): 57–77.

Klaiber, H. Allen and V. Kerry Smith. 20011a. "Quasi Experiments, Hedonic Models, and Estimating Tradeoffs for Local Amenities." *Working Paper*, August.

Klaiber, H. Allen and V. Kerry Smith. 2011b. "Preference Heterogeneity and Non-Market Benefits: The Roles of Structural Hedonic and Sorting Models." In Jeff Bennett, ed., *The International Yearbook of Non-Market Valuation* (Cheltenham, UK: Edward Elgar) pp. 222–253.

Kuminoff, Nicolai V. 2009. "Decomposing the Structural Identification of Nonmarket Values." *Journal of Environmental Economics and Management*, Vol. 57(2) March: 123–139.

Kuminoff, Nicolai V. and Jaren C. Pope. 2010. "Hedonic Equilibria, Land Value Capitalization, and the Willingness to Pay for Public Goods." May, *Working Paper.*

https://doi.org/10.1515/2152-2812.1083 Published online by Cambridge University Press

Kuminoff, Nicolai V., V. Kerry Smith and Christopher Timmins. 2010 "The New Economics of Equilibrium Sorting and its Transformational Role for Policy Evaluation." September, *NBER Working Paper* # 16349.

Nechyba, Thomas, J. 1999. "School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers." *Journal of Public Economic Theory*, Vol. 1(1) January: 5–50.

Nechyba, Thomas, J. 2000. "Mobility, Targeting, and Private-School Vouchers." *American Economic Review*, Vol. 90(1) March: 130–146.

Palmquist, Raymond B. 1992. "Valuing Localized Externalities." *Journal of Urban Economics*, Vol. 31(1) January: 59–68.

Parry, Ian, Roberton Williams and Lawrence Goulder. 1999. "When Can Carbon Abatement Policies Increase Welfare? The Fundamental Role of Distorted Factor Markets." *Journal of Environmental Economics and Management*, Vol. 37(1) January: 52–84.

Perroni, Carlo. 1992. "Homothetic Representations of Regular Non-Homothetic Preferences." *Economics Letters*, Vol. 40(1) September: 19–22.

Phaneuf, Daniel J., V. Kerry Smith, Raymond B. Palmquist and Jaren C. Pope, 2008, "Integrating Property Value and Local Recreation Models to Value Ecosystem Services in Urban Watersheds," *Land Economics*, Vol. 84(3) August: 361–381.

Poterba, James M. 1992. "Taxation and Housing: Old Questions, New Answers." *American Economic Review*, Vol. 82(2) May: 237–242.

Rutherford, Thomas F. 1997. "Applied General Equilibrium Modeling with MPSGE as a GAMS Subsystem: An Overview of the Modeling Framework and Syntax." *Unpublished Paper*, University of Colorado, Boulder, CO, March.

Rutherford, Thomas F. 2002. "Lecture Notes on Constant Elasticity Forms." *Unpublished Paper*, University of Colorado, Boulder, CO, November.

Scarf, Herbert. 1973. *The Computation of Economic Equilibrium* (New Haven: Yale University Press).

Shoven, John and John Whalley. 1992. *Applying General Equilibrium* (New York: Cambridge University Press).

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf and Randy Walsh. 2002. "Interjurisdictional Housing Prices in Location Equilibrium." *Journal of Urban Economics*, Vol. 52(1) July: 131–153.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf and Randy Walsh. 2004. "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods." *International Economic Review*, Vol. 45(4) November: 1047–1077.′

Smith, V. Kerry and Jared C. Carbone. 2007. "Should Benefit Cost Analyses Take Account of General Equilibrium Effects?" *Research in Law and Economics: A Journal of Policy*, Vol. 23 edited by Richard O. Zerbe Jr. (Amsterdam: Elsevier), pp. 247–272.

Smith, V. Kerry, Holger Sieg, H. Spencer Banzhaf and Randy Walsh. 2004. "General Equilibrium Benefits for Environmental Improvements: Projected Ozone Reductions under EPA's Prospective Analysis for the Los Angeles Air Basin." *Journal of Environmental Economics and Management*, Vol. 47(3) May: 559–584.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, Vol. 64(5) October: 416–424.

Walsh, Randall L. 2007. "Endogenous Open Space Amenities in a Locational Equilibrium." *Journal of Urban Economics*, Vol. 61(2) March: 319–344.

West, Sarah and Roberton Williams. 2004. "Estimates from a Consumer Demand System: Implications for the Incidence of Environmental Taxes." *Journal of Environmental Economics and Management*, Vol. 47(3) May: 535–558.

Whalley, John. 1975. "How Reliable is Partial Equilibrium Analysis?" *The Review of Economics and Statistics*, Vol. 57(3): 299-310

Whittington, Dale and Duncan MacRae Jr., 1986, "The Issue of Standing in Cost-Benefit Analysis," *Journal of Policy Analysis and Management*, Vol. 5(4): 665–682.

Willig, Robert. 1976. "Consumer's Surplus Without Apology." *American Economic Review*, Vol. 66(4) September: 589–597.