
Volume 185, Issue 1, 16 February 2008
ISSN 0377-2217
ELSEVIER
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
Available online at
ScienceDirect
www.sciencedirect.com
EURO
www.elsevier.com/locate/ejor

Available online at www.sciencedirect.com
ScienceDirect
European Journal of Operational Research 185 (2008) 195–203
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/ejor

Stochastics and Statistics

Comparative statics of the generalized maximum entropy estimator of the general linear model

Michael R. Caputo [a,*], Quirino Paris [b]

[a] Department of Economics, University of Central Florida, Orlando, FL 32816-1400, USA
[b] Department of Agricultural and Resource Economics, University of California, One Shields Avenue, Davis, CA 95616-8512, USA

Received 12 December 2005; accepted 16 December 2006

Abstract

The generalized maximum entropy method of information recovery requires that an analyst provides prior information in the form of finite bounds on the permissible values of the regression coefficients and error values for its implementation. Using a new development in the method of comparative statics, the sensitivity of the resulting coefficient and error estimates to the prior information is investigated. A negative semidefinite matrix reminiscent of the Slutsky-matrix of neoclassical microeconomic theory is shown to characterize the said sensitivity, and an upper bound for the rank of the matrix is derived.

Keywords: Comparative statics; Sensitivity analysis; Generalized maximum entropy; General linear model; Regression

1. Introduction

Alternative methods for the recovery of economic information show up routinely in the econometrics and statistics literature. Recently, a method of information recovery based on the maximum entropy formalism (Shannon, 1948; Jaynes, 1957) has been developed by Golan et al. (1996). Golan et al. (1996) developed what they call the generalized maximum entropy (GME) approach to information recovery, and related GME to ordinary least squares and other estimators using real data and Monte Carlo experiments. The basic idea underlying the GME approach is to reparameterize the general linear model $Y_t = \sum_{k=1}^{K} X_{tk}\beta_k + \varepsilon_t, t = 1, 2, \ldots, T$, so that it can be accommodated within the classical maximum entropy framework. This is accomplished by treating each regression coefficient β_k as a discrete random variable with a compact *support interval* consisting of $2 \leqslant M < +\infty$ possible outcomes, where a support interval is defined as a closed and bounded interval of the real line in which each β_k is restricted to lie. Taking the case of $M = 2$ for pedagogical reasons, and letting z_{k1} and z_{k2} be the two possible outcomes and thus the finite endpoints of the support interval, β_k is expressed

---

[*] Corresponding author. Tel.: +1 407 823 1405; fax: +1 407 823 3269.
E-mail address: mcaputo@bus.ucf.edu (M.R. Caputo).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2006.12.031

as a convex combination of z_{k1} and z_{k2}, namely, $\beta_k = p_1^k z_{k1} + p_2^k z_{k2}$, where $p_1^k \geqslant 0, p_2^k \geqslant 0$, and $p_1^k + p_2^k = 1$. Similarly, each error ε_t is treated as a finite and discrete random variable with a compact support interval consisting of $2 \leqslant N < +\infty$ possible outcomes. Assuming $N = 2$, each ε_t can also be written as a convex combination of the two finite support values v_{t1} and v_{t2}, to wit, $\varepsilon_t = q_1^t v_{t1} + q_2^t v_{t2}$, where $q_1^t \geqslant 0$, $q_2^t \geqslant 0$, and $q_1^t + q_2^t = 1$.

As a result of this reparameterization, the entropy objective function consists of the coefficient entropy plus the error entropy, with the reparameterized linear model and adding up conditions on the probabilities serving as constraints. Given the support intervals specified by the analyst, the solution of the GME problem recovers estimates of the probabilities $(\hat{p}_1^k, \hat{p}_2^k)$ and $(\hat{q}_1^t, \hat{q}_2^t)$ that maximize entropy and thus are the most uniform or uncertain. Accordingly, the GME solution recovers estimates of the probability distribution of each regression coefficient and error. Point estimates of the regression coefficients and errors are then obtained by using the formulas described above, videlicet, $\hat{\beta}_k = \hat{p}_1^k z_{k1} + \hat{p}_2^k z_{k2}$ and $\hat{\varepsilon}_t = \hat{q}_1^t v_{t1} + \hat{q}_2^t v_{t2}$.

Arguably, the most controversial aspect of the GME reparameterization is that the researcher specifies the support intervals of the regression coefficients and errors as priors. Given this feature, it is natural to ask the ensuing question—How sensitive are the estimated regression coefficients and errors to the prior information provided by the analyst? The central result of this paper is a precise analytical and qualitative answer to this question. In particular, we show that the change in the GME estimates of the regression coefficients and errors of the general linear statistical model with respect to a change in their support intervals—known as the comparative statics, or equivalently, sensitivity results, of the GME problem—are contained in a negative semidefinite matrix, and provide an upper bound to the rank of the matrix. Moreover, this matrix takes a form that is familiar to students of neoclassical microeconomic theory, scilicet, that of the Slutsky-matrix of the theory of the consumer, and explain why this is so.

In deriving the above results, we take the position of the "least informed researcher", or equivalently, what we refer to as the "worst case scenario", terms we shall rigorously define in Section 3. For the moment it is sufficient to say that we take the position of a researcher who has little or no a priori knowledge about the true values of the coefficients and errors to be estimated. Not only is this the archetypal situation in applied work, especially when working with flexible functional forms for which the individual coefficients typically have no economic meaning in and of themselves, but it is also the instance when specification of the support intervals is most crucial. We first show that in general, a primal comparative statics analysis of the GME problem does not yield refutable implications for the effects of changes in the coefficient and error support bounds on the individual coefficient and error estimates, and explain why this is the case.

In spite of the aforementioned lack of refutable sensitivity results in the GME problem, it is not the case that refutable qualitative implications are not forthcoming in the GME problem. It turns out that a primal view of the GME problem leads one down a path of logic that prevents the discovery of its intrinsic and refutable comparative statics properties. The new comparative statics formalism of Partovi and Caputo (2006) is thus employed to uncover the intrinsic semidefinite matrix that characterizes the refutable comparative statics properties of the GME problem.

2. Literature review

The Monte Carlo evidence presented by Golan et al. (1996) is quite compelling and typically shows the "superiority" of the GME coefficient estimates over that of OLS and other estimators in the case of ill-conditioned design matrices. This evidence, however, is not entirely convincing in our view seeing as the true values of the coefficients are known in Monte Carlo experiments, thus permitting the researcher to select the support intervals so that the true values of the coefficients are always contained in the support interval, and thereby resulting in GME coefficient estimates that are often "superior" to that of the other estimators. Applied economists, on the other hand, never know the true values of the coefficients they are attempting to estimate and therefore do not have such additional a priori information when specifying the support intervals for the coefficients and errors. Given that this is the universal situation in applied work, it is manifestly important to come to some general understanding regarding the sensitivity of the GME coefficient and error estimates to the specification of their support intervals. As has already been noted, this is precisely the issue we address in the paper.

Golan et al. (1997) presented Monte Carlo evidence on the sensitivity of a mean squared error (MSE) loss function with respect to changes in the coefficient and error supports for a Tobit specification. Based on the evidence presented in their Table 1(b), they concluded (p. 40) "…that, regardless of the upper and lower bounds of Z (the support interval), in all cases the GME is superior to the ML (maximum likelihood)". But in the same paragraph the authors supplied evidence that refutes this claim by specifying a narrow support interval that contains all the true coefficient values, but which turns out to have a MSE that is larger than all those reported in Table 1(b), including that of the ML estimator. In fact, the MSE varies by nearly 60% in their Monte Carlo experiments and is most sensitive when the support interval is narrow. Our view, therefore, is that the Monte Carlo evidence presented by Golan et al. (1997) illustrates the sensitivity of the coefficient estimates to changes in the support intervals.

Other researchers have also expressed a great deal of concern about the sensitivity of the GME coefficient estimates to the choice of their support intervals. For example, Léon et al. (1999) showed, via numerical sensitivity analysis, that their coefficient estimates were not invariant to the support intervals, and as a result concluded (p. 438) "…that great care should be given to the choice of appropriate support values". Similarly, Paris and Howitt (1998, p. 135) observed a "…large difference between the individual coefficients …" in response to different support intervals for the coefficients. Golan et al. (1996, p. 138) also recognized that the GME coefficient estimates are sensitive to the specification of their support intervals. They discussed this aspect of the GME formalism in two paragraphs in the context of a Monte Carlo experiment, though they did perform an analytical comparative statics or sensitivity exercise (p. 110) for the effect of the error support interval on the Lagrange multiplier in the Jaynes dice problem.

What is lacking in the GME literature, therefore, is a general, analytic, as well as complete discussion of the sensitivity of the GME coefficient and error estimates to the specification of their support intervals. Léon et al. (1999, Abstract) have also noted the lack of work on this crucial matter when they asserted that "…the sensitivity of the GME estimates with respect to the design of the prior information set needs to be investigated further". Our paper responds to these concerns and dearth of research by filling in this important gap in the GME literature.

3. Theoretical background of GME

Following Golan et al. (1996, Chapter 6), given a generic general linear model specification, say $Y_t = \sum_{k=1}^{K} X_{tk}\beta_k + \varepsilon_t, t = 1, 2, \ldots, T$, the GME nonlinear optimization problem is given by

$$\max_{\mathbf{p},\mathbf{q}} \quad E(\mathbf{p},\mathbf{q}) \overset{\text{def}}{=} -\sum_{k=1}^{K}\sum_{m=1}^{M} p_m^k \ln p_m^k - \sum_{t=1}^{T}\sum_{n=1}^{N} q_n^t \ln q_n^t \tag{1}$$

$$\begin{aligned}
\text{s.t.} \quad & Y_t = \sum_{k=1}^{K} X_{tk} \sum_{m=1}^{M} z_{km} p_m^k + \sum_{n=1}^{N} v_{tn} q_n^t, \quad t = 1, 2, \ldots, T, \\
& \sum_{m=1}^{M} p_m^k = 1, \ k = 1, 2, \ldots, K, \quad \sum_{n=1}^{N} q_n^t = 1, \ t = 1, 2, \ldots, T, \\
& p_m^k \geqslant 0, \ m = 1, 2, \ldots, M, \ k = 1, 2, \ldots, K, \quad q_n^t \geqslant 0, \ n = 1, 2, \ldots, N, \ t = 1, 2, \ldots, T,
\end{aligned}$$

where $\mathbf{p}^k \overset{\text{def}}{=} (p_1^k, p_2^k, \ldots, p_M^k), k = 1, 2, \ldots, K$, and $\mathbf{q}^t \overset{\text{def}}{=} (q_1^t, q_2^t, \ldots, q_N^t), t = 1, 2, \ldots, T$, are probability vectors with $\mathbf{p} \overset{\text{def}}{=} (\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^K)$ and $\mathbf{q} \overset{\text{def}}{=} (\mathbf{q}^1, \mathbf{q}^2, \ldots, \mathbf{q}^T), z_{km}, m = 1, 2, \ldots, M$, are the support values for the kth coefficient β_k, $k = 1, 2, \ldots, K$, thereby implying that $\beta_k = \sum_{m=1}^{M} z_{km} p_m^k, k = 1, 2, \ldots, K, v_{tn}, n = 1, 2, \ldots, N$, are the support values for the tth value of the error ε_t, $t = 1, 2, \ldots, T$, thus implying that $\varepsilon_t = \sum_{n=1}^{N} v_{tn} q_n^t, t = 1, 2, \ldots, T, Y_t, t = 1, 2, \ldots, T$, is the tth value of the dependent variable, and X_{tk}, $t = 1, 2, \ldots, T$, $k = 1, 2, \ldots, K$, is the tth value of the kth independent variable.

Problem (1) represents the GME framework for the general linear model at its highest level of generality, and as such results in unwieldy and unnecessarily complicated comparative statics expressions devoid of insight. Consequently, we turn to the case of the "least informed researcher" or "worst case scenario" first introduced in Section 1, as this is the case of importance in applied work. Recall that the "worst case scenario" was defined as a situation in which the researcher has little or no a priori knowledge about the true values of the coefficients and errors to be estimated. The key to implementing this point of view is to adopt a reparameterization of the regression coefficients and errors that formally reflects little a priori knowledge about them.

To implement the "worst case scenario" mathematically, assume that each coefficient and error has two support values, i.e., $M = N = 2$, thereby implying that

$$\beta_k = z_{k1}p_1^k + z_{k2}p_2^k, \quad k = 1,2,\ldots,K, \tag{2}$$

$$\varepsilon_t = v_{t1}q_1^t + v_{t2}q_2^t, \quad t = 1,2,\ldots,T. \tag{3}$$

Furthermore, in order to reflect that the analyst has essentially no a priori information about the true values of the coefficients and errors, assume that the support interval of each coefficient and error is symmetrically centered about the origin, that is to say, $z_{k1} = -z_{k2} < 0$, $k = 1,2,\ldots,K$, and $v_{t1} = -v_{t2} < 0$, $t = 1,2,\ldots,T$. Then upon defining the support values $z_k \overset{\text{def}}{=} z_{k2} > 0, k = 1,2,\ldots,K$, and $v_t \overset{\text{def}}{=} v_{t2} > 0, t = 1,2,\ldots,T$, Eqs. (2) and (3) can be rewritten as follows:

$$\beta_k = z_k[p_2^k - p_1^k], \quad k = 1,2,\ldots,K, \tag{4}$$

$$\varepsilon_t = v_t[q_2^t - q_1^t], \quad t = 1,2,\ldots,T. \tag{5}$$

Eqs. (4) and (5) constitute the reparameterization of the regression coefficients and errors in the case of the "least informed researcher" or "worst case scenario". In other words, in the "worst case scenario", each coefficient and error is reduced to a function of the symmetric endpoint of its support interval and the difference between the probability that the coefficient and error take on positive and negative values. The reparameterization in Eq. (4) therefore restricts the estimated value of the coefficient β_k to lie in the compact support interval $[-z_k, z_k]$. Even so, this does not change the interpretation of the regression coefficient β_k. This is because the interpretation of β_k is derived from the general linear model and the data used to estimate it. Similar remarks apply to the values of the errors in Eq. (5).

Using the "least informed researcher" reparameterization of the regression coefficients and errors given in Eqs. (4) and (5), we may rewrite the general GME problem (1) in the form

$$\begin{aligned}
\max_{\mathbf{p},\mathbf{q}} \quad & E(\mathbf{p},\mathbf{q}) \overset{\text{def}}{=} -\sum_{k=1}^{K}\sum_{m=1}^{2} p_m^k \ln p_m^k - \sum_{t=1}^{T}\sum_{n=1}^{2} q_n^t \ln q_n^t \\
\text{s.t.} \quad & Y_t = \sum_{k=1}^{K} X_{tk} z_k[p_2^k - p_1^k] + v_t[q_2^t - q_1^t], \quad t = 1,2,\ldots,T, \\
& \sum_{m=1}^{2} p_m^k = 1, \ \ k = 1,2,\ldots,K, \quad \sum_{n=1}^{2} q_n^t = 1, \ \ t = 1,2,\ldots,T.
\end{aligned} \tag{6}$$

Problem (6) is the mathematical representation of the GME problem from the point of view of the "least informed researcher", or that of the "worst case scenario". Observe that we have dropped the nonnegativity constraints on the probabilities in problem (6) seeing as they do not bind in the optimal solution.

In any given application of GME, the data Y_t, $t = 1,2,\ldots,T$, and X_{tk}, $t = 1,2,\ldots,T$, $k = 1,2,\ldots,K$, are given. Consequently, the parameters of interest for comparative statics or sensitivity purposes are the endpoints of the support intervals for the coefficients and errors, namely, $z_k > 0$, $k = 1,2,\ldots,K$, and $v_t > 0$, $t = 1,2,\ldots,T$, respectively, seeing as these are the subjective values supplied by the researcher. The comparative statics of interest are therefore, the effects of changes in the parameters $z_k > 0$, $k = 1,2,\ldots,K$, and $v_t > 0$, $t = 1,2,\ldots,$ T, on the probability vectors $\mathbf{p}$ and $\mathbf{q}$, the latter two of which are the *decision variables* of the GME problem. Also of interest are the sensitivities of the GME estimates of the coefficient vector $\boldsymbol{\beta}$ and error vector $\boldsymbol{\epsilon}$.

For the "worst case scenario" GME problem given in Eq. (6), define the optimal values of the probabilities as $\hat{\mathbf{p}}^k(\mathbf{a}) \overset{\text{def}}{=} (\hat{p}_1^k(\mathbf{a}), \hat{p}_2^k(\mathbf{a})), k = 1,2,\ldots,K$, and $\hat{\mathbf{q}}^t(\mathbf{a}) \overset{\text{def}}{=} (\hat{q}_1^t(\mathbf{a}), \hat{q}_2^t(\mathbf{a})), t = 1,2,\ldots,T$, and the corresponding optimal values of the regression coefficients and errors as $\hat{\beta}_k(\mathbf{a}) = z_k[\hat{p}_2^k(\mathbf{a}) - \hat{p}_1^k(\mathbf{a})], k = 1,2,\ldots,K$, and $\hat{\varepsilon}_t(\mathbf{a}) = v_t[\hat{q}_2^t(\mathbf{a}) - \hat{q}_1^t(\mathbf{a})], t = 1,2,\ldots,T$, respectively, where $\mathbf{a} \overset{\text{def}}{=} (z_1, z_2, \ldots, z_K, v_1, v_2, \ldots, v_T) \in \mathbb{R}_{++}^{K+T}$ is the parameter vector used in the sensitivity analysis.

Having stated the GME problem of interest and defined the variables and parameters to be used in the comparative statics analysis, we close this section by pointing out a general feature of the "worst case scenario" GME problem (6), to wit, that there are no refutable comparative statics results, in general, for the effects of changes in the coefficient and error supports on the individual probabilities, coefficients, and errors. In other words, one cannot generally sign the partial derivative comparative statics of the form $\partial \hat{p}_2^k(\mathbf{a})/\partial z_1$ for problem (6). This is a result of the fact noted by Silberberg (1978, p. 293) that when a parameter appears in a constraint of an optimization problem, unambiguous signs for the partial derivative comparative statics will not be forthcoming in general for that parameter, and the fact that *all* the parameters of problem (6) enter every general linear model constraint.

4. Refutable comparative statics

Our objective in this section is to derive a semidefinite matrix that is comprised of the partial derivatives of the optimal values of the decision variables with respect to the parameters for the "worst case scenario" GME problem (6). Before doing so, we provide a brief exposition of Theorem 1 of Partovi and Caputo (2006), the main result used in our sensitivity calculations.

The key idea of the comparative statics method of Partovi and Caputo (2006) originates in the observation that the partial derivatives of the decision variables with respect to the parameters, considered individually, are not necessarily susceptible to a refutable comparative statics characterization in a given optimization problem. This observation is manifest in the archetype utility maximization problem, where a linear combination of such partial derivatives, namely, those with respect to price and income, is required for an exhaustive comparative statics characterization of the problem as embodied in the Slutsky-matrix. On the other hand, aside from an inessential scale factor, a general, linear combination of partial derivatives with respect to the parameters is just a directional derivative pointed in a particular direction in parameter space. However, not every arbitrary direction in parameter space will suffice, and it turns out that the constraint structure of the optimization problem plays a crucial role in determining the required directions in parameter space if the desired semidefiniteness property is to emerge with the constraints already implemented. As shown in Partovi and Caputo (2006, Lemma 1), the key to such constraint-free comparative statics results is to find directional derivatives with respect to the parameters that return zero when applied to the constraint functions of the optimization problem. Directional derivatives with this null property are defined as *generalized compensated derivatives* (GCD's) by Partovi and Caputo (2006).

Using the above null requirement, one can characterize the desired directions in parameter space in simple geometrical terms: all tangential directions with respect to the level set of the constraint function in parameter space will yield directional derivatives with the null property. Equivalently, all directions on the tangent hyperplane to the level surface of the constraint function in parameter space have the requisite property. Because the gradient vector of a function is orthogonal to the level set of the function, any vector that lies in the null space of the gradient of the function with respect to the parameters lies in the tangent hyperplane to the level set of that function in parameter space. Hence the desired directions in parameter space are those orthogonal to the gradient of the constraint function with respect to the parameters. How many of these directions and corresponding GCD's are necessary or desirable? In general, this number equals the dimension of the tangent hyperplane. Accordingly, a set of GCD's constructed from the basis vectors of the tangent hyperplane provides a *complete* set of GCD's (Partovi and Caputo, 2006). Using this simple prescription, we will next construct a complete set of GCD's and use it in conjunction with Theorem 1 of Partovi and Caputo (2006) to derive a complete comparative statics characterization of the "worst case scenario" GME problem (6).

In order to develop a better understanding of the construction of the complete set of GCD's for problem (6), we pause momentarily and devote the next three paragraphs to a detailed discussion of their construction in a simple, low-dimensional setting. After this preliminary material is finished, we return to the case under investigation in this section, videlicet, the "worst case scenario" GME problem (6).

To begin the detailed construction of the complete set of GCD's in the simple, low-dimensional case, assume that $K=2$ and $T=1$. These assumptions imply a general linear model of the form $Y_1 = X_{11}z_1[p_2^1 - p_1^1] + X_{12}z_2[p_2^2 - p_1^2] + v_1[q_2^1 - q_1^1]$, which, in this instance, is the only constraint in problem (6) that contains the regression coefficient and error support values. Defining $g(\mathbf{x};\mathbf{a}) \overset{\text{def}}{=} Y_1 - X_{11}z_1[p_2^1 - p_1^1] - X_{12}z_2[p_2^2 - p_1^2] - v_1[q_2^1 - q_1^1]$ as the value of the constraint function and $\mathbf{x} \stackrel{\text{def}}{=} (p_1^1, p_2^1, p_1^2, p_2^2, q_1^1, q_2^1)$ as the decision vector, its gradient vector with respect to the parameters $\mathbf{a} \stackrel{\text{def}}{=} (z_1, z_2, v_1)$ is given by

$$\nabla^{\mathbf{a}} g(\mathbf{x};\mathbf{a}) = (-X_{11}[p_2^1 - p_1^1], -X_{12}[p_2^2 - p_1^2], -[q_2^1 - q_1^1]).$$

Under the present simplifying assumptions the dimension of the parameter space is three and the dimension of the space spanned by the gradient vector $\nabla^{\mathbf{a}} g(\mathbf{x};\mathbf{a})$ is one as long as at least one probability is not equal to one-half. Hence, the dimension of the space spanned by the tangent hyperplane to the constraint surface is two. This means that two vectors are required for a complete description of the tangent hyperplane to the constraint surface in parameter space. We therefore seek two vectors that lie in the null space of the gradient vector, i.e., two vectors, say $\mathbf{t}^{\alpha} \stackrel{\text{def}}{=} (t_1^{\alpha}, t_2^{\alpha}, t_3^{\alpha})$, $\alpha = 1, 2$, that satisfy $\mathbf{t}^{\alpha} \cdot \nabla^{\mathbf{a}} g(\mathbf{x};\mathbf{a}) = 0$ for $\alpha = 1, 2$, and which are linearly independent. Such vectors are defined as *isovectors* by Partovi and Caputo (2006). We are therefore led to the following system of linear equations for the isovectors:

$$\begin{bmatrix} -X_{11}[p_2^1 - p_1^1] & -X_{12}[p_2^2 - p_1^2] & -[q_2^1 - q_1^1] \end{bmatrix} \begin{bmatrix} t_1^1 & t_1^2 \\ t_2^1 & t_2^2 \\ t_3^1 & t_3^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

For each isovector $\mathbf{t}^{\alpha} \stackrel{\text{def}}{=} (t_1^{\alpha}, t_2^{\alpha}, t_3^{\alpha})$, $\alpha = 1, 2$, we have one equation and three unknowns, resulting in two degrees of freedom for the choice of each isovector, and thus infinitely many choices for each isovector. But this is exactly what one would expect given that there are infinitely many basis vectors for a two-dimensional real vector space.

The "art" in deriving the isovectors is in determining the most economically informative ones from the infinite set of choices. Our experience using this method of comparative statics indicates that the "cleanest" set of isovectors is typically the best in the sense that it permits the simplest and richest economic interpretation of the resulting comparative statics matrix. We follow this prescription and so begin the construction of the isovectors by setting $t_1^1 = 1$ and $t_2^1 = 0$, seeing as we noted above that two degrees of freedom exist for choosing any given isovector. Substituting the choices $t_1^1 = 1$ and $t_2^1 = 0$ into the first of the above linear equation yields $t_3^1 = -X_{11}[p_2^1 - p_1^1]/[q_2^1 - q_1^1]$. Similarly, for the second isovector set $t_1^2 = 0$ and $t_2^2 = 1$, so that upon solving the second linear equation above we get $t_3^2 = -X_{12}[p_2^2 - p_1^2]/[q_2^1 - q_1^1]$. The two isovectors are therefore given by

$$\mathbf{t}^1 \stackrel{\text{def}}{=} \left(1, 0, \frac{-X_{11}[p_2^1 - p_1^1]}{[q_2^1 - q_1^1]}\right) = \left(1, 0, -X_{11}\frac{[\beta_1/z_1]}{[\varepsilon_1/v_1]}\right),$$
$$\mathbf{t}^2 \stackrel{\text{def}}{=} \left(0, 1, \frac{-X_{12}[p_2^2 - p_1^2]}{[q_2^1 - q_1^1]}\right) = \left(0, 1, -X_{12}\frac{[\beta_2/z_2]}{[\varepsilon_1/v_1]}\right),$$

the second equality for each isovector resulting from the equations $\beta_k/z_k = [p_2^k - p_1^k], k = 1, 2,$ and $\varepsilon_1/v_1 = [q_2^1 - q_1^1]$. These isovectors are linearly independent and span the tangent hyperplane to the level curve of the constraint function in view of the fact that the only solution to the linear system $c_1\mathbf{t}^1 + c_2\mathbf{t}^2 = \mathbf{0}_3$ is $c_1 = c_2 = 0$. Thus the pair of isovectors $(\mathbf{t}^1, \mathbf{t}^2)$ form a basis for the tangent hyperplane to the constraint surface in the (z_1, z_2, v_1)-space.

Turning to the corresponding complete set of GCD's, scilicet $D_{\alpha}(\mathbf{x};\mathbf{a})$, $\alpha = 1, 2$, they are by definition constructed according to the rule $D_{\alpha}(\mathbf{x};\mathbf{a}) \stackrel{\text{def}}{=} \mathbf{t}^{\alpha} \cdot \nabla^{\mathbf{a}}, \alpha = 1, 2$. For the simplified case presently under discussion, we have

$$D_1(\mathbf{x};\mathbf{a}) \stackrel{\text{def}}{=} \mathbf{t}^1 \cdot \nabla^{\mathbf{a}} = \left(1, 0, -X_{11}\frac{[\beta_1/z_1]}{[\varepsilon_1/v_1]}\right) \cdot \left(\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}, \frac{\partial}{\partial v_1}\right) = \frac{\partial}{\partial z_1} - X_{11}\frac{[\beta_1/z_1]}{[\varepsilon_1/v_1]}\frac{\partial}{\partial v_1},$$
$$D_2(\mathbf{x};\mathbf{a}) \stackrel{\text{def}}{=} \mathbf{t}^2 \cdot \nabla^{\mathbf{a}} = \left(0, 1, -X_{12}\frac{[\beta_2/z_2]}{[\varepsilon_1/v_1]}\right) \cdot \left(\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}, \frac{\partial}{\partial v_1}\right) = \frac{\partial}{\partial z_2} - X_{12}\frac{[\beta_2/z_2]}{[\varepsilon_1/v_1]}\frac{\partial}{\partial v_1}.$$

These GCD's are clearly of the flavor of the Slutsky compensated derivative of neoclassical consumer theory. For example, the partial derivative $\frac{\partial}{\partial z_1}$ is analogous to the price derivative in the Slutsky compensated derivative, the weight $-X_{11}[\beta_1/z_1]/[\varepsilon_1/v_1]$ is analogous to the commodity weight on the income derivative, and the partial derivative $\frac{\partial}{\partial v_1}$ is analogous to the income derivative. Verification that these are in fact GCD's is straightforward:

$$D_1(\mathbf{x};\mathbf{a})\circ g(\mathbf{x};\mathbf{a}) = \left[\frac{\partial}{\partial z_1} - X_{11}\frac{[\beta_1/z_1]}{[\varepsilon_1/v_1]}\frac{\partial}{\partial v_1}\right]\circ\left[Y_1 - X_{11}z_1[p_2^1-p_1^1] - X_{12}z_2[p_2^2-p_1^2] - v_1[q_2^1-q_1^1]\right]$$

$$= -X_{11}[p_2^1-p_1^1] - X_{11}\frac{[\beta_1/z_1]}{[\varepsilon_1/v_1]}\left[-[q_2^1-q_1^1]\right] = -X_{11}[p_2^1-p_1^1] + X_{11}[\beta_1/z_1] = 0,$$

$$D_2(\mathbf{x};\mathbf{a})\circ g(\mathbf{x};\mathbf{a}) = \left[\frac{\partial}{\partial z_2} - X_{12}\frac{[\beta_2/z_2]}{[\varepsilon_1/v_1]}\frac{\partial}{\partial v_1}\right]\circ\left[Y_1 - X_{11}z_1[p_2^1-p_1^1] - X_{12}z_2[p_2^2-p_1^2] - v_1[q_2^1-q_1^1]\right]$$

$$= -X_{12}[p_2^2-p_1^2] - X_{12}\frac{[\beta_2/z_2]}{[\varepsilon_1/v_1]}\left[-[q_2^1-q_1^1]\right] = -X_{12}[p_2^2-p_1^2] + X_{12}[\beta_2/z_2] = 0,$$

as $\beta_k/z_k = [p_2^k - p_1^k], k = 1,2$, and $\varepsilon_1/v_1 = [q_2^1 - q_1^1]$. Having given a rather detailed discussion of the construction of the isovectors and GCD's under the simplifying assumptions $K = 2$ and $T = 1$, we return to the general case of the "least informed researcher" in Eq. (6).

To begin, recall that the constraints of problem (6) that are functions of **a** are given by

$$g^t(\mathbf{x};\mathbf{a}) \overset{\text{def}}{=} Y_t - \sum_{k=1}^{K} X_{tk}z_k[p_2^k - p_1^k] - v_t[q_2^t - q_1^t], \quad t = 1,2,\ldots,T, \tag{7}$$

where $\mathbf{a} \overset{\text{def}}{=} (z_1, z_2, \ldots, z_K, v_1, v_2, \ldots, v_T) \in \mathbb{R}^{K+T}_{++}$ is the parameter vector defined in Section 3 and $\mathbf{x} \overset{\text{def}}{=} (\mathbf{p},\mathbf{q})$ is the decision vector. The gradient of $g^t(\cdot)$ with respect to **a** is then given by

$$\nabla^{\mathbf{a}} g^t(\mathbf{x};\mathbf{a}) = \left(-X_{t1}[p_2^1-p_1^1], \ldots, -X_{tK}[p_2^K-p_1^K], 0_1, \ldots, 0_{t-1}, -[q_2^t-q_1^t], 0_{t+1}, \ldots, 0_T\right) \in \mathbb{R}^{K+T}, \tag{8}$$

for $t = 1,2,\ldots,T$. Because the dimension of the parameter space is $K + T$ and the dimension of the space spanned by the gradient vectors $\nabla^{\mathbf{a}} g^t(\mathbf{x};\mathbf{a})$ is T for $q_2^t \neq q_1^t$, $t = 1,2,\ldots,T$, the dimension of the space spanned by the tangent hyperplane to the constraint surface is therefore K by the implicit function theorem. This means that K vectors are required for a complete description of the tangent hyperplane to the constraint surface in parameter space.

By showing that $\mathbf{t}^\alpha \cdot \nabla^{\mathbf{a}} g^t(\mathbf{x};\mathbf{a}) = 0$ for $\alpha = 1,2,\ldots,K$ and $t = 1,2,\ldots,T$, one can verify that the ensuing K vectors lie in the tangent hyperplane to the constraint surface in parameter space:

$$\mathbf{t}^\alpha \overset{\text{def}}{=} \left(\mathbf{e}_K^\alpha, \frac{-X_{1\alpha}[p_2^\alpha - p_1^\alpha]}{[q_2^1 - q_1^1]}, \frac{-X_{2\alpha}[p_2^\alpha - p_1^\alpha]}{[q_2^2 - q_1^2]}, \ldots, \frac{-X_{T\alpha}[p_2^\alpha - p_1^\alpha]}{[q_2^T - q_1^T]}\right) \in \mathbb{R}^{K+T}, \quad \alpha = 1,2,\ldots,K, \tag{9}$$

where $\mathbf{e}_K^\alpha \overset{\text{def}}{=} (0_1, 0_2, \ldots, 0_{\alpha-1}, 1_\alpha, 0_{\alpha+1}, \ldots, 0_K)$ is the standard basis vector in $\mathbb{R}^K$. That the above K vectors form a basis for the tangent hyperplane to the constraint manifold in parameter space and are thus isovectors, can be seen by constructing a $(K + T) \times K$ matrix whose columns are the isovectors $\mathbf{t}^\alpha$, and then noting that the rank of said matrix is K. Therefore, by definition, the complete set of GCD's is constructed by taking the Euclidean inner product of the isovectors $\mathbf{t}^\alpha$ given in Eq. (9) with the gradient operator $\nabla^{\mathbf{a}} \overset{\text{def}}{=} (\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}, \ldots, \frac{\partial}{\partial z_K}, \frac{\partial}{\partial v_1}, \frac{\partial}{\partial v_2}, \ldots, \frac{\partial}{\partial v_T})$, thereby yielding

$$D_\alpha(\mathbf{x};\mathbf{a}) \overset{\text{def}}{=} \mathbf{t}^\alpha \cdot \nabla^{\mathbf{a}} = \frac{\partial}{\partial z_\alpha} - \left[\frac{\beta_\alpha}{z_\alpha}\right] \sum_{s=1}^{T} \frac{X_{s\alpha}}{[\varepsilon_s/v_s]}\frac{\partial}{\partial v_s}, \quad \alpha = 1,2,\ldots,K, \tag{10}$$

where we have made use of $\beta_k/z_k = [p_2^k - p_1^k]$, $k = 1,2,\ldots,K$, from Eq. (4) and $\varepsilon_t/v_t = [q_2^t - q_1^t]$, $t = 1,2,\ldots,T$, from Eq. (5). The K directional derivatives $D_\alpha(\mathbf{x};\mathbf{a})$ given in Eq. (10) form a complete set of GCD's because $D_\alpha(\mathbf{x};\mathbf{a})\circ g^t(\mathbf{x};\mathbf{a}) = 0$ for $\alpha = 1,2,\ldots,K$ and $t = 1,2,\ldots,T$. With the construction of the complete set of GCD's for problem (6) finished, we turn to the last detail in preparation for the statement of our main result.

The Lagrangian function $L(\cdot)$ for problem (6) is defined as

$$L(\mathbf{p},\mathbf{q},\boldsymbol{\lambda},\boldsymbol{\pi},\boldsymbol{\theta};\mathbf{a}) \overset{\text{def}}{=} -\sum_{k=1}^{K}\sum_{m=1}^{2} p_m^k \ln p_m^k - \sum_{t=1}^{T}\sum_{n=1}^{2} q_n^t \ln q_n^t + \sum_{t=1}^{T}\lambda_t\left[Y_t - \sum_{k=1}^{K} X_{tk} z_k [p_2^k - p_1^k] - v_t[q_2^t - q_1^t]\right] + \sum_{k=1}^{K}\pi_k\left[1 - \sum_{m=1}^{2} p_m^k\right] + \sum_{t=1}^{T}\theta_t\left[1 - \sum_{n=1}^{2} q_n^t\right],$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{\mathrm{T}}$ is the Lagrange multiplier vector for the general linear model constraints, $\boldsymbol{\pi} \in \mathbb{R}^K$ is the Lagrange multiplier vector for the coefficient probability constraints, and $\boldsymbol{\theta} \in \mathbb{R}^{\mathrm{T}}$ is the Lagrange multiplier vector for the error probability constraints. The first-order necessary conditions of problem (6) include

$$\frac{\partial L}{\partial p_1^i} = -1 - \ln p_1^i + \sum_{t=1}^{T}\lambda_t X_{ti} z_i - \pi_i = 0, \quad i = 1,2,\ldots,K,$$

$$\frac{\partial L}{\partial p_2^i} = -1 - \ln p_2^i - \sum_{t=1}^{T}\lambda_t X_{ti} z_i - \pi_i = 0, \quad i = 1,2,\ldots,K,$$

$$\frac{\partial L}{\partial q_1^j} = -1 - \ln q_1^j + \lambda_j v_j - \theta_j = 0, \quad j = 1,2,\ldots,T,$$

$$\frac{\partial L}{\partial q_2^j} = -1 - \ln q_2^j - \lambda_j v_j - \theta_j = 0, \quad j = 1,2,\ldots,T,$$

along with the three sets of constraints. The ensuing theorem is the central result of our paper. Its proof follows from a straightforward application of Theorems 1 and 4 of Partovi and Caputo (2006) to problem (6), using the GCD's in Eq. (10) and the above first-order necessary conditions of problem (6).

Theorem 1 (Complete comparative statics). *The* $K \times K$ *comparative statics matrix* $\Psi(\mathbf{a})$ *for the GME problem (6) is symmetric and negative semidefinite, where*

$$\Psi_{\alpha\gamma}(\mathbf{a}) \overset{\text{def}}{=} \left[\sum_{t=1}^{T}\hat{\lambda}_t(\mathbf{a})X_{t\alpha}\right]\left[\frac{\partial[\hat{\beta}_\alpha(\mathbf{a})/z_\alpha]}{\partial z_\gamma} - \left[\frac{\hat{\beta}_\gamma(\mathbf{a})}{z_\gamma}\right]\sum_{s=1}^{T}\frac{X_{s\gamma}}{[\hat{\varepsilon}_s(\mathbf{a})/v_s]}\frac{\partial[\hat{\beta}_\alpha(\mathbf{a})/z_\alpha]}{\partial v_s}\right] - \left[\frac{\hat{\beta}_\alpha(\mathbf{a})}{z_\alpha}\right]\sum_{j=1}^{T}\frac{\hat{\lambda}_j(\mathbf{a})X_{j\alpha}}{[\hat{\varepsilon}_j(\mathbf{a})/v_j]}\left[\frac{\partial[\hat{\varepsilon}_j(\mathbf{a})/v_j]}{\partial z_\gamma} - \left[\frac{\hat{\beta}_\gamma(\mathbf{a})}{z_\gamma}\right]\sum_{s=1}^{T}\frac{X_{s\gamma}}{[\hat{\varepsilon}_s(\mathbf{a})/v_s]}\frac{\partial[\hat{\varepsilon}_j(\mathbf{a})/v_j]}{\partial v_s}\right],$$

$\alpha,\gamma = 1,2,\ldots,K$. *Moreover, the rank of* $\Psi(\mathbf{a})$ *is no larger than K.*

Theorem 1 contains all of the qualitative information derivable from problem (6) without imposing additional assumptions on its structure. The comparative statics matrix $\Psi(\mathbf{a})$ shows that it is the compensated changes in the support values that results in refutable comparative statics or sensitivity properties for problem (6). It is important to note that the form of the comparative statics given by Theorem 1 applies not to the coefficients and errors, but to their values relative to the endpoint of their support interval.

The comparative statics matrix $\Psi(\mathbf{a})$ consists of a linear combination of $T+1$ Slutsky-like forms, the latter consisting of a coefficient support effect and a linear combination of T error support effects. The resemblance of $\Psi(\mathbf{a})$ to the Slutsky-matrix of neoclassical consumer theory is not unexpected. This is because (i) only the decision variables enter the objective function of problem (6), and (ii) all of the coefficient and error support values enter its constraints linearly, just like in the prototype utility maximization problem. That there are T error support effects is explained by the fact that when the support interval for the αth coefficient changes, i.e., when z_α changes, the support intervals for all T errors must change in order for the compensation to be implemented so that the direction of change remains on the tangent hyperplane to the constraint manifold in parameter space. This makes sense too, seeing as a change in the support interval of any given coefficient affects all the general linear model constraints of problem (6).

In closing out this section, we return to the simplifying assumptions $K = 2$ and $T = 1$ in order to better understand Theorem 1. In this instance the typical element of $\Psi(\mathbf{a})$ is of the form

$$\begin{aligned}\Psi_{\alpha\gamma}(\mathbf{a}) \overset{\text{def}}{=} [\hat{\lambda}_1(\mathbf{a})X_{1\alpha}] \Bigg\{ & \left[\frac{\partial[\hat{\beta}_\alpha(\mathbf{a})/z_\alpha]}{\partial z_\gamma} - X_{1\gamma} \left[\frac{[\hat{\beta}_\gamma(\mathbf{a})/z_\gamma]}{[\hat{\varepsilon}_1(\mathbf{a})/v_1]} \right] \frac{\partial[\hat{\beta}_\alpha(\mathbf{a})/z_\alpha]}{\partial v_1} \right] \\ & - \left[\frac{[\hat{\beta}_\alpha(\mathbf{a})/z_\alpha]}{[\hat{\varepsilon}_1(\mathbf{a})/v_1]} \right] \left[\frac{\partial[\hat{\varepsilon}_1(\mathbf{a})/v_1]}{\partial z_\gamma} - X_{1\gamma} \left[\frac{[\hat{\beta}_\gamma(\mathbf{a})/z_\gamma]}{[\hat{\varepsilon}_1(\mathbf{a})/v_1]} \right] \frac{\partial[\hat{\varepsilon}_1(\mathbf{a})/v_1]}{\partial v_1} \right] \Bigg\}, \quad \alpha, \gamma = 1, 2.\end{aligned}$$

This result can be derived from Theorem 1, or by applying Theorem 1 of Partovi and Caputo (2006) to the first-order necessary conditions using the GCD's $D_1(\mathbf{x};\mathbf{a})$ and $D_2(\mathbf{x};\mathbf{a})$ and the equations $\beta_k/z_k = [2p_2^k - 1], k = 1, 2$, and $\varepsilon_1/v_1 = [2q_2^1 - 1]$. The Slutsky-like nature of the comparative statics matrix $\Psi(\mathbf{a})$ is self-evident. It consists of a linear combination of two Slutsky-like terms under the simplifying assumptions $K = 2$ and $T = 1$. The form of $\Psi(\mathbf{a})$ shows that even in this special case it is not possible to derive refutable comparative statics results for the effects of the support values on the individual coefficients and errors, just as recognized in Section 3.

5. Summary and conclusion

It is indisputable that GME coefficient and error estimates are sensitive to the specification of their support intervals. What was heretofore not understood was whether intrinsic and refutable comparative statics results existed, and the form that such comparative statics would take. Our paper has provided a rather complete answer and assessment of these issues. Though in general there are no refutable partial derivative-type comparative statics results for the "worst case scenario" GME problem (6), the use of a new comparative statics formalism led to the discovery of a negative semidefinite matrix that provides a complete qualitative characterization of its refutable comparative statics properties. Moreover, the said matrix turns out to be similar in flavor to the ubiquitous Slutsky-matrix of neoclassical consumer theory, albeit in a generalized form, and it shows that it is linear combinations of the ratios of the coefficients and errors to the endpoints of their support intervals that posses refutable comparative statics properties.

Acknowledgements

Quirino Paris is a member of the Giannini Foundation of Agricultural Economics. We thank Richard E. Howitt, Martine Quinzii, and seminar participants at the Economic Theory Seminar Series at University of California, Davis for insightful remarks, and two anonymous reviewers for comments on the organization and presentation of the contents of the paper.

References

Golan, A., Judge, G., Miller, D., 1996. Maximum Entropy Econometrics. Wiley, Chichester UK.

Golan, A., Judge, G., Perloff, J., 1997. Estimation and inference with censored and ordered multinomial response data. Journal of Econometrics 79, 23–51.

Jaynes, E.T., 1957. Information theory and statistical mechanics. Physics Review 106, 620–630.

Léon, Y., Peeters, L., Quinqu, M., Surry, Y., 1999. The use of maximum entropy to estimate input–output coefficients from regional farm accounting data. Journal of Agricultural Economics 50, 425–439.

Paris, Q., Howitt, R.E., 1998. An analysis of ill-posed production problems using maximum entropy. American Journal of Agricultural Economics 80, 124–138.

Partovi, M.H., Caputo, M.R., 2006. A complete theory of comparative statics for differentiable optimization problems. Metroeconomica 57, 31–67.

Shannon, C.E., 1948. A mathematical theory of communication. The Bell System Technical Journal 27, 379–423.

Silberberg, E., 1978. The Structure of Economics: A Mathematical Analysis. McGraw-Hill Publishing Company, New York.