

Estimating Parameters and Structural Change in CGE Models Using a Bayesian Cross-Entropy Estimation Approach

Delfin S. Go, Hans Lofgren, Fabian Mendez Ramos, and Sherman Robinson¹

Draft: June 13, 2014

Not for quotation

Abstract

This paper uses a two-step Bayesian cross-entropy estimation approach in an environment of noisy and scarce data to estimate behavioral parameters for a computable general equilibrium (CGE) model and to measure how labor augmenting productivity and/or other parameters in the model shift over time to generate historically observed changes in economic structure. In this approach, the parameters in a CGE model are treated as fixed but unknown, which we represent as prior mean values with prior error mass functions. Estimation of (and inference about) the parameters involves using an information-theoretic Bayesian approach to exploit additional information in the form of new data from a series of Social Accounting Matrices (SAMs), which we assume were measured with error. The estimation procedure is “efficient” in the sense that it uses all available information and makes no assumptions about unavailable information. As illustrations, we apply the methodology to estimate the parameters of a CGE model using data for South Korea and for Sub-Saharan Africa.

JEL codes: C68, C61, C11, C13, E16, J24

Keywords: Computable general equilibrium (CGE) models, Bayesian estimation, cross-entropy estimation

¹ Delfin S. Go – The World Bank; Hans Lofgren – The World Bank; Fabian Mendez Ramos- The World Bank; and Sherman Robinson – International Food Policy Research Institute. The study was partially funded by a grant from the Knowledge for Change Program II Trust Fund of the World Bank. The views expressed are those of the authors and do not necessarily reflect those of their respective institutions or affiliated organizations. We thank Syud Amer Ahmed, Maryla Maliszewska, Karen Thierfelder, Hans Timmer, Dominique van der Mensbrugghe, and Dirk Willenbockel for several discussions, comments and suggestions at various points of the study.

Introduction

CGE models are often criticized for weak empirical estimation of the parameters and for preserving static economic structures in long-term simulations. The problem is associated with the lack of reliable time-series data in developing countries to support standard econometric estimation of parameters and shifts in structure. In particular, we have lacked a times series of social accounting matrices (SAMs) and their associated prices and quantities that provide the information base for these models. There are examples of econometric estimation of parts of the CGE model where some time series data are available—see Jorgenson and Yun (2013) and other work by Jorgenson and various coauthors (such as Jorgenson, 2011, Jorgenson and Timmer 2011, Jorgenson et al. 2012).

This paper uses an information theoretic cross-entropy estimation approach in an environment of scarce data measured with error to estimate behavioral parameters, labor augmenting productivity, or other parameters in the model that shift over time to generate historically observed changes in economic structure. With noisy and limited observations common to developing countries, the unknown parameters of a CGE model cannot generally be observed and measured directly. In systems and information theory, this is described as a *stochastic inverse problem*: how to use available “information” to recover these unknown and uncontrolled components. Robust solution methods consistent with the underlying ill-posed noisy information recovery problem have been developed under Bayesian inference and decision theory (see, for example, Golan et al. 1996 and Judge and Mittelhammer 2012).² In CGE model estimation, Arndt et al. (2002) used the cross entropy estimation method from information theory to estimate the trade substitution elasticities in a CGE model of Mozambique. Robinson et al. (2001) also utilized the cross-entropy method in data updating and estimation when limited information is available to construct a balanced new SAM.

In this paper, we extend and refine the cross entropy estimation method to account for different levels of noise and amount of information in a sequential and particular way. We have started with a collection of SAMs with identical accounts for the Republic of Korea (South Korea) for various years, which is a major advance on available data to support estimation, although still not enough to support standard econometric methods. Each new data set for later years provides new information to improve the estimation of parameters of a CGE model, updating priors that initially were based on scattered data and theoretical properties of various functions in the model. The statistical results recovered from this noisy and limited informational environment are clearly dependent on how the CGE model is formulated; the more data, the

² Not always initially accepted, the Bayesian approach is now applied widely in decision and information theory, operations research, and more recently, in finance and macroeconomics, such as the recent methodology regarding dynamic stochastic general equilibrium (DSGE) modeling. A popular history of its development, the debates about its validity and uses, its triumphant applications in the great wars of the last century, and its recent uses is available in McGrayne, S. B. (2011).

more reliable the estimates of structural parameters. The estimation procedure is “efficient” in the sense that it uses all available information and makes no assumptions about unavailable information. Importantly,³ estimated parameters maintain consistency with the microeconomic foundations of the general equilibrium theory embodied in CGE models.

We make use of the collection of available comparable SAMs for different years to estimate parameters and to account for equilibrium dynamics of the country’s economic flows as postulated in the CGE model. Since SAMs are expressed in nominal values and information about relative prices may be limited, we also implement a data step by using the cross-entropy method to estimate SAMs consistent with available information on changes in prices before parameter estimation. We also examine whether convergence and stability of the estimated parameters is evident when adding additional SAMs. We consider various discrete prior probability distributions, specifying additive or multiplicative errors on the behavioral and structural parameters and SAM targets to examine convergence and stability results. While the SAMs only provide nominal flows, we use scattered information on relative prices and factor accumulation quantities, where available. Since developing countries are undergoing fundamental economic transformations as they grow, our method can allow for evolution of a number of parameters, including the elasticities, factor specific technical change, or shifts in value added and trade shares that may be important in the changing economic structure. As more estimates are made for many countries, any regularity in the estimates may further inform the pattern of development and structural change, following the work of Chenery and various coauthors (Chenery et al. 1975, Chenery and Elkington 1979, Chenery et al. 1986).

In the case of South Korea, the CGE/SAM/cross-entropy (CGE-SAM-CE) method is applied to the period 1990 to 2011 to estimate economic behavioral parameters such as the elasticities of substitution between traded goods and domestic goods and between factors in production functions. We also explore for possible evolution in economic structure as reflected in the growth of total factor productivity (TFP), labor-augmenting productivity in the value added function of each sector, or in trends of trade elasticities and trade shares over time.

As datasets for global CGE modeling are increasingly compiled by efforts such as the Global Trade Analysis Project (GTAP) at Purdue University, regional units are becoming popular in CGE modeling, but the availability of regional relative price data still lags. We therefore also look at a regional application to Sub-Saharan Africa, where data noise and constraints can be severe in order to demonstrate how the approach may be applied.

As a last point, the framework as implemented permits both the recovery of behavioral (unobserved) parameters in the estimation model and a switch to policy simulations in the estimated CGE application mode once the parameters are estimated.

³ See Arnold Zellner’s definition of an efficient “information processing rule” in Zellner (1988).

The rest of the paper is organized as follows. Section II presents the methodology where we describe the two-step, cross entropy approach employed in the paper for estimating the behavioral parameters and shifts in the economic structure. Using the CGE-SAM-CE method Section III discusses the estimates and results as well as the error measures corresponding to the estimates. In the last section, we conclude and suggest areas for future research.

Methodology

Inferring information about parameters and elements of a CGE model of a developing country is often made difficult because of incomplete, infrequent, or uncertain evidence such as dated or poorly constructed SAMs and auxiliary data. There could be other related reasons. Time and budget constraints may prevent the difficult and tedious task of gathering the needed data for better estimation and calibration. For students and practitioners studying CGE modeling for the first time, having a preliminary model running for learning purposes is usually a first step in the decision whether to invest seriously in a more appropriate model and its associated data.

To deal with the severe data constraints, some practical procedures were commonly applied: subjective and expert judgments about the parameters and other assumptions; borrowing estimates from other studies; and sensitivity testing to refine their values and the economic reasonableness of simulation results. With somewhat more information, “back-casting” or “double-calibration” procedures were employed to improve the validation of models (e.g., Dawkins et al. 2001, Okusima and Tamura 2009). Although Bayesian in spirit, these procedures do not assess the inherent errors or noise associated with the weak data or with the estimates recovered in the process. To deal with the noisy and limited data more directly, the method of cross-entropy (CE) estimation from information theory and Bayesian econometrics has recently been deployed to update and rebalance a SAM or to separately estimate related parameters in the CGE model (Robinson et al. 2001, and Arndt et al. 2002).⁴ In that approach, initial estimates are refined with new information so that they maximize the probability of matching the historical record of vital aggregates in a SAM and they minimize a statistically measured distance of the calibration to the unknown parameters.

In this paper, we extend the cross-entropy method by integrating the SAM data preparation with the CGE parameter estimation in a two-step procedure: i) a data step that adjusts the historical SAMs to a common base year taking into account that the SAMs are measured with errors and relative price indices are scant; and ii) a parameter estimation step that calculates (filters) parameters and structural change simultaneously within the specifications of the CGE model. We follow the customary calibration process of a CGE model in order to

⁴ There are links between cross-entropy and empirical-likelihood estimation procedures. For example, with lots of data they will converge to the same distribution. See Golan et al. (1996) and Judge and Mittelhammer (2012).

provide a practical approach and to ensure the consistency with economic theory in the behavioral equations. Finally, we develop a reference programming code in GAMS for possible wider use of the technique that matches CGE variables straight to the SAM cells, which in turn greatly simplifies the structure of the code.⁵ In what follows, we describe briefly the different elements of the methodology and how it is applied to estimate and calibrate a CGE model.

The data step

As the basic source of information for a CGE model, a SAM is a starting point in the calibration process. Although a time-series of SAMs is still scarce for many countries, more and more SAMs are being constructed by national statistical agencies. For example, they are being compiled and reconciled in globally consistent data sets by efforts such as GTAP. As more SAMs are estimated, each additional SAM provides new information to combine with previous SAMs to test and adjust prior assumptions about the parameters and components in the CGE model.

In this data step, we exploit a basic property that each SAM is a transaction (flow) matrix approach to national accounting (see Stone 1962, and Pyatt and Round 1985). Robinson et al. (2001) originally employ the cross-entropy estimation method, using available limited data such as macro aggregates, in order to update a SAM when a new complete SAM is not available. Here, we apply the methodology for the purpose of pooling and adjusting all available historical SAMs when auxiliary data to directly do so are incomplete, essentially making use of the SAM's necessary consistency with macroeconomic aggregates in the national accounts.

More precisely, the first step is to adjust and scale the nominal SAMs associated with different years so that they are measured in the same base year units and are hence comparable. The ideal way to do this is to separate the nominal magnitudes into their respective prices and quantities; and each cell in the SAM should be adjusted by the correct relative price index so that *all* individual cells are expressed in equivalent real terms. Where that is possible, it should clearly be employed. For many developing countries, however, the informational requirements will be daunting relative to the state of their statistical capacities (e.g. countries in Sub-Saharan Africa or newly formed countries after civil conflicts). Noise in the data, including the quality of price indices if available and the values of the specific SAM cells, will likely be very high.

In the absence of full information about prices and quantities, each SAM still brings new information about the economic structure of the country, albeit in nominal values. It adds new observational data and the new evidence can be used to update the old priors about the

⁵ GAMS refers to the General Algebraic Programming System of GAMS Development Corporation. The GAMS code in this paper is available from the authors in the future when documentation is completed.

parameters. Mindful of the data noise and limitations however, we implement the following practical steps:

- a. First, deflate all cells in the historical SAMS by a single overall GDP deflator so that they are broadly expressed in the same base year prices as a first step. For models with regional aggregation of countries, this may be the only means possible if deriving weighted aggregation of country specific prices is not feasible.
- b. We assign greater weight to known data in a limited informational environment. Accordingly, as a second step (after deflating the SAMs), we apply the cross entropy estimation method to target known macroeconomic aggregates in constant prices from the national accounts, which are generally available for most countries; and we rebalance the SAM to preserve, given priors on measurement error, these critical economic aggregates. This approach in effect extends the SAM cross-entropy estimation method to a constant price application.
- c. Alternatively, one could use known relative price indices (relative to the GDP deflator) to target the macroeconomic aggregates. The latter could conceivably be expanded information-wise if auxiliary data such as world prices of exports and imports for various commodities are also available. As mentioned, the more cells in the SAMs that can be deflated to the right base year prices, the better the economic representation that is reflected in the SAM. In any case, to ensure that the SAM adheres to known national account aggregates in constant prices, it should still be estimated and rebalanced appropriately.
- d. To avoid potential scaling problems when the years of the SAMs are far apart, it is possible to further scale the SAMS to a similar GDP level using the GDP of a particular year or the average of all years.
- e. In the extreme case, where there are no price and national accounts data to work with (e.g. in post-conflict or data-poor economies), the evolution of the economic structure reflected in the nominal SAMs still adds information to test various prior assumptions about the unknown parameters. It is still better than using less information from a single SAM.

The method used here follows the earlier approach in Robinson et al. (2001) in that standard additive errors and discrete probability distributions are defined at three levels of information – (1) each cell value in the SAM; (2) the column totals in the SAM (SAM balance condition); and (3) the macroeconomic aggregates in the SAM. The formal mathematical

statement generally follows the earlier work and is not repeated here. In adjusting the historical SAMs to constant prices under limited information, we assign the smallest standard errors to the macro aggregates since they are known from the national accounts.

The parameter estimation step

In this section, we describe the cross-entropy method to estimate the unknown parameters and other components of the economic framework, subject to the SAM data, their aggregation and adjustments (the cross-entropy method in the first or SAM preparation and estimation stage) and the specifications of the CGE model.

Prior and posterior probability distributions of parameter error. Our basic approach explicitly recognizes that the value of each parameter θ_i is unknown, but can be represented by its prior θ_i^0 and an error term ε_{θ_i} , in additive or multiplicative way, i.e., $\theta_i = \theta_i^0 + \varepsilon_{\theta_i}$ or $\ln(\theta_i) = \ln(\theta_i^0) + \varepsilon_{\theta_i}$, respectively. Moreover, the error term has an assigned prior error probability mass distribution, such that $Prob(\varepsilon_{\theta_i} \in E_{j,\theta_i} : \theta_i \subseteq E_{j,\theta_i}) = \pi(\varepsilon_{j,\theta_i}) \in [0,1]$, $\varepsilon_{\theta_i} \in \{E_{1,\theta_i}, E_{2,\theta_i}, \dots, E_{n,\theta_i}\}$, $Prob(E_{j,\theta_i} \cap E_{k \neq j,\theta_i}) = 0$ and $\sum_j \pi(\varepsilon_{j,\theta_i}) = 1$. Then, our information theoretic cross-entropy estimation centers on the posterior parameter values, $\Theta = f(F(\cdot) = \mathbf{0}, \mathbf{Y}, \mathbf{V}, \Pi(\boldsymbol{\varepsilon} | \cdot))$, conditioned on the specifications of the CGE model $F(\dots) = \mathbf{0}$, the SAM targets, $\mathbf{Y}$, the SAM data, $\mathbf{V}$, and the posterior error distribution $\Pi(\varepsilon_{\theta_1}, \dots, \varepsilon_{\theta_I} | \mathbf{Y}, \mathbf{V}, F(\dots) = \mathbf{0})$, which is a conditional joint distribution of all the error parameters, $\boldsymbol{\varepsilon}(\varepsilon_{\theta_1}, \dots, \varepsilon_{\theta_I})$.

Mathematical statement of the problem. Two concepts, precision and prediction, play important roles in the interpretation of the method. “Precision” refers to the behavioral and structural parameters and to the difference between the posterior and prior values of these parameters. “Prediction” refers to the sample data and the difference between the estimated values of the targeted (selected) SAM cells and their prior values.

We set up the CGE-CE method as a minimization problem of an objective function (1) subject to equations (2) through (13) as its constraints (Appendix 2 lists various notations and their definitions). More specifically, the cross entropy objective function minimizes the sum of the cross-entropy deviation of the prior and estimated probability weight, for all the discrete error distributions that characterize the unobserved parameters and SAM targets:

$$(1) \quad \min_{\{\mathbf{w}, \mathbf{x}_t\}} \left\{ \alpha_1 \sum_{m=1}^M \sum_{k_m=1}^{K_m} w_{B,k_m} \ln \left(\frac{w_{B,k_m}}{\bar{w}_{B,k_m}} \right) + \alpha_2 \sum_{t=1}^T \sum_{s=1}^S \sum_{k_s=1}^{K_s} w_{Z,t,k_s} \ln \left(\frac{w_{Z,t,k_s}}{\bar{w}_{Z,t,k_s}} \right) \right. \\ \left. + \alpha_3 \sum_{t=1}^T \sum_{n=1}^N \sum_{k_n=1}^{K_n} w_{Y,t,k_n} \ln \left(\frac{w_{Y,t,k_n}}{\bar{w}_{Y,t,k_n}} \right) \right\}$$

subject to:⁶

The CGE block:

$$(2) \quad \mathbf{F}(\mathbf{X}_t, \mathbf{Z}_t^o, \mathbf{Z}_t^u, \mathbf{B}, \boldsymbol{\delta}) = \mathbf{0}, \quad \forall t \in T.$$

The calibration block:

$$(3) \quad \boldsymbol{\delta} = \Phi(\mathbf{Z}_{t_0}, \mathbf{B}).$$

The behavioral parameter (precision) block:

$$(4) \quad B_m = B_m^0 e^{e_{B,m}}, \quad \forall m \in M.$$

$$(5) \quad e_{B,m} = \sum_{k_m=1}^{K_m} w_{B,k_m} v_{B,k_m}, \quad \forall m \in M.$$

$$(6) \quad \sum_{k_m=1}^{K_m} w_{B,k_m} = 1, \quad \forall m \in M.$$

The unobserved exogenous or non-behavioral parameter (precision) block:

$$(7) \quad \mathbf{Z}_t^u = \mathbf{Z}_t^{u,0} e^{e_{Z,t}}, \quad \forall t \in T.$$

$$(8) \quad e_{Z,t,s} = \sum_{k_s=1}^{K_s} w_{Z,t,k_s} v_{Z,t,k_s}, \quad \forall s \in S.$$

$$(9) \quad \sum_{k_s=1}^{K_s} w_{Z,t,k_s} = 1, \quad \forall s \in S.$$

The SAMs target (prediction) block:

$$(10) \quad \mathbf{Y}_t = \mathbf{Y}_t^0 e^{e_{Y,t}}, \quad \forall t \in T.$$

$$(11) \quad \mathbf{V}_t = \mathbf{G}(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{Z}_t^o, \mathbf{Z}_t^u, \mathbf{B}, \boldsymbol{\delta}), \quad \forall t \in T.$$

$$(12) \quad e_{Y,t,n} = \sum_{k_n=1}^{K_n} w_{Y,t,k_n} v_{Y,t,k_n}, \quad \forall n \in N, \forall t \in T.$$

$$(13) \quad \sum_{k_n=1}^{K_n} w_{Y,t,k_n} = 1, \quad \forall n \in N, \forall t \in T.$$

The first set of constraints belongs to the standard CGE block of equations as presented in (2). This is the representation of the general equilibrium model, a square system of non-linear equations that satisfy the property of homogeneity of degree zero in prices. This $\mathbf{F}(\cdot)$ vector function depends on the CGE endogenous variables $\mathbf{X}_t$, the observed exogenous variables $\mathbf{Z}_t^o$, the unobserved exogenous variables $\mathbf{Z}_t^u$, the behavioral parameters $\mathbf{B}$, and the calibration restrictions $\boldsymbol{\delta}$. The index t denote the set of time observations $\{t_1, t_2, \dots, t_T\}$ s.t. $t \in \mathbb{N}_{>1}$.

⁶ The vector of posterior error probabilities ($\mathbf{w}$) is partitioned in three components: the behavioral parameter, the SAM target and the unobserved (exogenous) parameters, i.e., $\mathbf{w} = \{\mathbf{w}_B, \mathbf{w}_Y, \mathbf{w}_Z\}$.

The next constraints belong to the calibration block (3). This set of equations depends on the vector of exogenous variables $\mathbf{Z}$ at the base period t_0 and on the behavioral parameters $\mathbf{B}$. Note that $t_0 \subset t$, which means that this base period can be selected from any element of the sequence corresponding to SAM's periods.

The behavioral parameter block is given by equation (4) to (6). The entropy setup requires modeling the unobserved (behavioral) parameters $\mathbf{B}$ by using prior information $\mathbf{B}^0$ and an error term $\mathbf{e}_B$. In equation (4), the error term is entered in multiplicative form, such that $\mathbf{e}_B$ measures some amount of error between the prior $\mathbf{B}^0$ and posterior $\mathbf{B}$ information in logarithmic units. However, this can also be written in additive manner, i.e., $\mathbf{B} = \mathbf{B}^0 + \mathbf{e}_B$, then $\mathbf{e}_B$ will be in the same units as the behavioral parameters. Equation (5) models the error terms $e_{B,m}$ as an expected value using error support values v_{B,k_m} and their respective endogenously estimated probability weights w_{B,k_m} . This behavioral parameter block requires that these discrete error probability weights sum to one (6). The prior specified for the probability weights w_{B,k_m} in the estimation reflect the degree of information available about the distributions—how “informative” are the priors.⁷

Additionally, in this minimization problem the restrictions presented in equations (7) through (9), pertains to the unobserved exogenous (non-behavioral) parameters $\mathbf{Z}_t^u$. $\mathbf{Z}_t^u$ depends on prior information $\mathbf{Z}_t^{u,0}$ and their error terms $e^{e_{Z,t}}$ in a multiplicative way as presented in (7). The unobserved exogenous parameters $\mathbf{Z}_t^u$ may contain parameters such as technical and productivity coefficients and rates. As in the two previous blocks, equation (8) also restricts the error terms with the expected value formula, considering specific support values and particular posterior weights, v_{Z,t,k_s} and w_{Z,t,k_s} , respectively. The corresponding discrete error probabilities are required to sum to one, equation (9).

Finally, the last set of constraints defines the SAMs target block through equations (10) to (13). The SAM targets are selected cells of the SAMs that serve to adjust the system for an easier convergence and a closer representation of the historical data. One of the main differences between Arndt et al. (2002) and our method is that our targets are specific SAM cells (flows in constant currency units), while the model of 2002 was targeting a subset of the endogenous CGE variables ($\mathbf{X}_t$). Note that this method permits targeting specific cells of the SAM at particular periods, for instance, household savings could be a SAM target variable for period t_1 but not for time t_T . The first equation in this block, (10), models the SAM targets $\mathbf{Y}_t$ as function of their prior values $\mathbf{Y}_t^0$ and their multiplicative error terms $e^{e_{Y,t}}$. Additionally, equation (11) is the representation of the endogenous recovery of the complete SAMs $\mathbf{V}_t$. That is, the full posterior SAMs are recovered as functions of the endogenous CGE variables $\mathbf{X}_t$, the SAM targets $\mathbf{Y}_t$, the observed and unobserved exogenous variables, $\mathbf{Z}_t^o$ and $\mathbf{Z}_t^u$, respectively, the behavioral

⁷ The specification of “uninformative” and “informative” priors is described in detail in Appendix 3.

parameters $\mathbf{B}$, and the calibration constraints δ . Equation (12) also imposes the expected value restriction on the error terms of the SAM targets using error supports v_{Y,t,k_n} and endogenously estimated posterior weights w_{Y,t,k_n} . The last set of constraints of this SAM target block belongs to equation (13) which again requires that the discrete error probability weights sum to one.

Considering that the selection of error types and prior distributions for the SAM targets and unobserved parameters are fixed, then the number of possible combinations of SAM targets depends on the number of non-empty cells in each SAM and the number of available SAM periods. Thus, the maximum number of combinations is $2^{(N_{t_1}^{SAM} + N_{t_2}^{SAM} + \dots + N_{t_T}^{SAM})}$, where $N_{t_1}^{SAM}$ represents the total number of non-empty cells in the SAM of period t_1 , and the subindex T stands for the last period of the available SAMs. Furthermore, the possibility to update new SAM observations plus the error term prior distributions with old optimal posteriors will make this procedure a Bayesian estimation process of successive prior-to-posterior-to prior, etc. steps.

An alternative way to express the objective function is using expected values for the error probabilities as in (14), with expected values defined as $\mathbb{E}_{w_B}[\ln(w_{B,k_m})] = \sum_{k_m=1}^{K_m} w_{B,k_m} \ln(w_{B,k_m})$ and $\mathbb{E}_{w_B}[\ln(\bar{w}_{B,k_m})] = \sum_{k_m=1}^{K_m} w_{B,k_m} \ln(\bar{w}_{B,k_m})$. In this formulation, $\mathbf{w}$ and $\bar{\mathbf{w}}$ represent the matrices with the posterior and prior error probability mass function values of the correspondent error elements. Each of the expected value sub-indexes expresses their own and specific probability spaces. Under the afore-mentioned definitions, we can explain the value of the objective function as a noise measure of the observed data, or more precisely, as a pseudo-distance⁸ of expected data noise. In this case, the bigger the value of the objective function, the noisier the data, but the more informative the parameters. Our optimization problem therefore measures and minimizes the expected pseudo-measure of error probabilities or data noise.

$$\begin{aligned}
 (14) \quad \min_{\{\mathbf{w}, \mathbf{X}_t\}} & \{ \alpha_1 \sum_{m=1}^M (\mathbb{E}_{w_B}[\ln(w_{B,k_m})] - \mathbb{E}_{w_B}[\ln(\bar{w}_{B,k_m})]) \\
 & \min_{\{\mathbf{w}, \mathbf{X}_t\}} \left\{ \alpha_1 \sum_{m=1}^M (\mathbb{E}_{w_B}[\ln(w_{B,k_m})] - \mathbb{E}_{w_B}[\ln(\bar{w}_{B,k_m})]) \right. \\
 & + \alpha_2 \sum_{t=1}^T \sum_{s=1}^S (\mathbb{E}_{w_{Z,t}}[\ln(w_{Z,t,k_s})] - \mathbb{E}_{w_{Z,t}}[\ln(\bar{w}_{Z,t,k_s})]) \\
 & \left. + \alpha_3 \sum_{t=1}^T \sum_{n=1}^N (\mathbb{E}_{w_{Y,t}}[\ln(w_{Y,t,k_n})] - \mathbb{E}_{w_{Y,t}}[\ln(\bar{w}_{Y,t,k_n})]) \right\}.
 \end{aligned}$$

As introduced above, equation (1) includes the precision and prediction parts; the precision part is represented by every term that is multiplied by α_1 and α_2 , while α_3 corresponds to the section related with the “prediction” part embedded on the SAM target errors. Thus, at the optimal levels, the weights α_1 , α_2 and α_3 partially define the noise level of this cross entropy

⁸ See Judge, G. G. and R. C. Mittelhammer (2012, p. 107) for a brief discussion.

objective function, and the more weight on any of the above alpha parameters, the bigger marginal values of the correspondent error probabilities.

Inference and Goodness of Fit. In making inference about the recovered parameters and error terms, each of the elements of $\mathbf{B}$, $\mathbf{Z}_t^u$, and SAM targeted cells $\mathbf{Y}_t$ can be generalized and expressed in the alternative forms: $\theta = \theta^0 + \varepsilon$ or $\ln \theta = \ln \theta^0 + \ln \varepsilon$, where ε is the stochastic error part with probability weights $\mathbf{w}$, prior weights $\bar{\mathbf{w}}$ and support $\mathbf{v}$. After minimizing our CGE-CE objective function, the recovered error terms $\hat{\varepsilon}$ measure how good the CGE system and cross entropy priors $\mathbf{B}^o$, $\mathbf{Z}_t^{u,o}$, and $\mathbf{Y}_t^o$ are capturing the (information) noise. On this regard, we are interested to test the behavior of the estimated error terms, $\hat{\varepsilon}$.

Applying a goodness-of-fit statistic for the error probability weights, we test the null hypothesis $H_0: \mathbf{w} = \bar{\mathbf{w}}$ versus the alternative $H_1: \mathbf{w} \neq \bar{\mathbf{w}}$ with significance level α (e.g. $\alpha=0.05$) for a specific parameter error. Then the test statistic in (15) would reject the null if $\hat{\chi}^2 > \chi_{(\alpha, K-1)}^2$, where K is the number of support elements in the error (discrete) distribution.

$$(15) \quad \hat{\chi}_{(K-1)}^2 = \sum_{k=1}^K \frac{(w_k - \bar{w}_k)^2}{\bar{w}_k}$$

$$\hat{\chi}_{(K-1)}^2 = \sum_{k=1}^K \frac{(w_k - \bar{w}_k)^2}{\bar{w}_k}$$

Moreover, in order to summarize the CGE-CE results, we also generate block tests grouping the probability error terms of: 1) the beta (behavioral) parameters; 2) the unobserved (technical) parameters; 3) the SAM targets; and 4) the sum of the three aforementioned blocks. Assuming independence among the correspondent error terms $\varepsilon_{B,m}$, $\varepsilon_{Y,t,n}$ and $\varepsilon_{Z,t,s}$, we estimate the block goodness-of-fit tests as presented in (16), which is the statistic for the Beta parameter errors.

$$(16) \quad \hat{\chi}_{([\sum_{m=1}^M k_m] - M)}^2 = \sum_{m=1}^M \sum_{k_m=1}^{K_m} \frac{(w_{B,k_m} - \bar{w}_{B,k_m})^2}{\bar{w}_{B,k_m}}$$

$$\hat{\chi}_{([\sum_{m=1}^M k_m] - M)}^2 = \sum_{m=1}^M \sum_{k_m=1}^{K_m} \frac{(w_{B,k_m} - \bar{w}_{B,k_m})^2}{\bar{w}_{B,k_m}}$$

Additionally, the statistics in (15) and (16) can have a second order approximation using the Kullback-Liebler (cross entropy) distance function as presented in (17).

$$(17) \quad D(\mathbf{w} \parallel \bar{\mathbf{w}}) \equiv \sum_{k=1}^K w_k \ln \left(\frac{w_k}{\bar{w}_k} \right) \cong \frac{1}{2} \sum_{k=1}^K \frac{(w_k - \bar{w}_k)^2}{\bar{w}_k}$$

$$D(\mathbf{w} \parallel \bar{\mathbf{w}}) \equiv \sum_{k=1}^K w_k \ln \left(\frac{w_k}{\bar{w}_k} \right) \cong \frac{1}{2} \sum_{k=1}^K \frac{(w_k - \bar{w}_k)^2}{\bar{w}_k}$$

For individual and subgroup parameter estimates, we compute model pseudo $-R^2$ statistics (18), where $\hat{S}(\mathbf{w})$ stands for the normalized entropy of the error terms. Since $\hat{S}(\mathbf{w}) \in [0,1]$, a value $\hat{S}(\mathbf{w}) = 0$ implies no uncertainty while $\hat{S}(\mathbf{w}) = 1$ redirects total uncertainty in the sense that $\mathbf{w}$ is uniformly distributed. Thus, in this CGE-SAM-CE setup as pointed out in $\theta = \mathcal{Y}\theta^0 + \varepsilon$ (or $\ln \theta = \ln \theta^0 + \ln \varepsilon$), $\hat{R}^2 = 0$ means that the CGE-SAM-CE system has total uncertainty about the error behavior of θ , while $\hat{R}^2 = 1$ reveals that CGE-CE setup has total certainty on the prior error behavior of θ .

$$(18) \quad \hat{R}^2 = 1 - \hat{S}(\mathbf{w}) = 1 - (-\sum_{k=1}^K w_k \ln w_k) / \ln K.$$

Finally, a third type of statistic is presented in (19). In this case, $\tilde{S}(\mathbf{w}) \in [0, \infty^+)$, thus, a value close to 1 means that the system has a proper prior selection of the error distribution (probability weights) while something different than one reflects the noise inconsistencies between our initial prior and the posterior error values (we over/under predict the parameter error behavior).

$$(19) \quad \tilde{S}(\mathbf{w}) = (-\sum_{k=1}^K w_k \ln w_k) / (-\sum_{k=1}^K \bar{w}_k \ln \bar{w}_k).$$

Estimation and Results

The cross entropy estimation method is technically compatible with any CGE model. In the implementation, we choose a widely used standard, the IFPRI model (Löfgren, H., et al. 2002) as recently implemented by (Cicowiez, M. and H. Lofgren 2006). This is a static CGE model for a single open economy in the tradition of Dervis, K., et al. (1982) and De Melo, J. and S. Robinson (1989). Over the years, CGE models of this type have been applied to a wide range of analysis of economic policy and external shocks in developing countries; hence, it is a good starting point for illustrations (see a recent survey of CGE models with policy applications to developing countries by Devarajan and Robinson 2013). Since the specifications of the standard CGE model is well documented, we only briefly sketch its outline to emphasize the key unknown parameters of behavioral relationships for estimation. Appendix 1 figure depicts the general structure of the production and consumption sides in the model, respectively.

SAM preparation

For illustrations of the methodology, we select South Korea as a country case and Sub-Saharan Africa as a regional application.

Although there is good and extensive data available in the case of South Korea, we do not include all of them in order to demonstrate parameter estimation in a more limited data circumstances that are likely to prevail in other developing countries. Accordingly, we make use

5 specific SAMs of South Korea for the following years, 1990, 1995, 2000, 2005, and 2011. We aggregate the SAMs to cover 6 sectors – agriculture, mining, manufacturing, utilities, construction, and services. As described in the data step, we assign the following standard errors for the three target level of information in each SAM: 0.3333 for the cells; 0.3333 for the column totals; and 0.01 for the macro aggregates. Table 1a illustrates how the macro results in percent deviations from the numbers of the national accounts.

Table 1b shows a similar application at a regional level, the case of Sub-Saharan Africa, using the regional SAMs for 2004, 2007, and 2011 from the GTAP database. The GTAP SAMs are already expressed in U.S. dollar, which is convenient for the data step since regional income accounts are available in the World Development Indicators at the World Bank in both current and constant price series. This means relevant price indices are derivable for regional GDP and for several of its components. The data step in this case consists of deflating the SAM values by first using the U.S GDP deflator and then targeting macroeconomic aggregates that are available.⁹ Regional accounts were available for the GDP components of consumption as well as value added figures for mining, utilities, and construction; hence they were not targeted in the data step for SSA. The GTAP SAM for SSA in 2011 is also preliminary and the deviations from the regional accounts targets are higher than those in the other two years.¹⁰ Even so, they still provide valuable observations to use, albeit more noisy, in the Bayesian entropy approach.

Table 1a: Percent Deviations of Estimated Versus Control Values of National Accounts, The Case of South Korea

(Based on different SAM years, in constant 2005 prices)

	1990	1995	2000	2005	2011
Personal consumption	-1.807	-2.849	-0.173	0.135	0.601
Government	-0.395	-0.551	-0.071	0.028	0.118
Investment	-0.772	-2.101	-0.065	0.078	0.228
Exports	0.057	0.048	0.053	0.007	0.005
Agriculture value added	0.049	0.043	0.022	-0.015	-0.036
Mining value added	0.031	0.038	0.027	0.001	-0.024
Manufacturing value added	0.052	0.058	0.027	-0.017	-0.041
Utilities value added	0.042	0.035	0.032	0.001	-0.038
Construction value added	0.024	0.004	0.011	-0.009	-0.028
Value added for all activities	2.597	2.987	0.330	-0.226	-1.397

Source: Authors' calculations.

⁹ The regional macroeconomic aggregates are available from World Bank World Development Indicators.

¹⁰ Given bigger divergences between the 2010 GTAP SAM information and the World Bank World Development Indicators (WDI) macro data and to achieve an optimal convergence in the 2010 SSA SAM, we had to allow a macro target standard error of 0.05 (instead of the 0.01 as before pointed out).

Table 1b: Percent Deviations of Estimated Versus Control Values of National Accounts, The Case of Sub-Saharan Africa
(Based on different SAM years, in constant 2004 prices)

	2004	2007	2011
Government	0.018	-0.014	-2.192
Investment	0.001	-0.086	-10.315
Exports	0.715	0.674	-14.999
Imports	-0.563	-0.908	-11.215
Agriculture value added	0.044	-0.083	-6.952
Manufacturing value added	0.082	0.013	-0.443
Services value added	0.234	-0.119	-9.493
Value added of all activities	-0.544	0.103	-6.064

Source: Authors' calculations.

Elasticities

Elasticities are critical unknown behavioral parameters in the standard CGE model, which are usually defined by using the CES functions (Arrow et al. 1961). For example, the trade elasticities define the substitution possibilities between domestic and foreign goods in the CET (constant elasticity of transformation) and CES (constant elasticity of substitution) functions. Suppressing the sector and time subscripts, the two functions can be written symmetrically, using the same form,

$$X = \bar{A} \left[\sum \delta_i \cdot (\lambda_i \cdot x_i)^\rho \right]^{\frac{1}{\rho}}$$

In the two factor case, X is the CES or CET composite of factor x_1 and x_2 , $\bar{A}$ the shift parameter, δ_i the CES share or distribution parameter and $\delta_1 + \delta_2 = 1$, λ_i a factor augmenting or biased productivity parameter for x_i , and ρ the exponent: $X = F(x_1, x_2; \delta_1, \delta_2, \rho, \bar{A}, \lambda_1, \lambda_2)$. The CES substitution elasticity σ and CET transformation elasticity Ω are given by $\sigma = 1/(1 - \rho)$; $-\infty < \rho < 1$ in the CES case and $\Omega = 1/(\rho - 1)$; $1 < \rho < \infty$ in the CET case.

Below is the first order condition of the CES case, which is expressed in value terms of factor input per unit of output:

$$\frac{w_i \cdot x_i}{P \cdot X} = \bar{A}^{\sigma-1} \cdot \lambda_i^{\sigma-1} \cdot \delta_i^\sigma \left[\frac{P}{w_i} \right]^{\sigma-1}$$

where P is the price of the composite good X and w_i is the price of the input x_i . Alternatively, the value ratio of factor inputs for a homothetic aggregation function is the familiar:

$$\frac{w_1 \cdot x_1}{w_2 \cdot x_2} = \left[\frac{\lambda_2}{\lambda_1} \right]^{1-\sigma} \cdot \left[\frac{\delta_1}{\delta_2} \right]^\sigma \cdot \left[\frac{w_2}{w_1} \right]^{\sigma-1}$$

If there are only two components in the CES function, an index may summarize the relative factor-augmenting productivity, that is $\lambda = \lambda_2/\lambda_1$. Since both the CET and CES functions exhibit constant returns to scale, the allocation of the composite good depends only on the relative prices of the individual components.

When data problems prevent traditional econometric estimation of the elasticities, they are often assumed or taken from other studies as extraneous estimates. The normal CGE calibration commonly uses a single and most recent SAM. Accordingly, prices in the model are initialized to 1.0, and values for the parameters, $\bar{A}$ and δ , are then derived based on the information in the SAM and the assumed values of the elasticities. However, if there are older historical SAMs, the priors can be improved with our approach and it is a mistake to discard the older historical SAMs, which contain vital information to improve extraneous numbers.

Furthermore, unless the prior estimates are initially close to the unknown values of the parameters, the cross-entropy method will mostly improve the priors with additional SAMs; and the revised estimates can deviate significantly from the initial guesses. Even a single price deflator will generally bring about improved estimates over the priors. It is easy to see why from the two first-order conditions that are expressed in value shares above.

In the absence of good price indices, the right-hand side clearly contains an error by a factor of $\left[\frac{p}{w_i} \right]^{\sigma-1}$ or $\left[\frac{w_2}{w_1} \right]^{\sigma-1}$ if nominal values of the SAMs are used in the left-hand side of the equations for the implicit derivation of the elasticities. Clearly, if relative price changes (between the SAMs) are stable, the error will be small; otherwise, they can be significant. Even so, the Bayesian method will make use of whatever implicit price indices are available and allows for errors and prior probability to be assigned and the posterior probability to be computed from the cross-entropy likelihood objective function. In the SAM estimation step, we give more weight or certainty to known national accounts aggregates or values that are deflated by appropriate prices. Likewise, in the CGE estimation and calibration our method can also give more weight to a more recent SAM or to a particular SAM that is constructed with greater reliability. These steps will almost certainly improve the prior estimates.

Furthermore, our Bayesian method estimates all the parameters in the CGE model simultaneously relative to their initial or prior values, subject to the specifications of the CGE model and the information from all available SAMs. Table 2a and b show the results of the cross-entropy estimation of the Armington elasticities for South Korea and SSA, respectively

under 2 cases. Case 1 assumes that only the GDP deflator is available to adjust the SAMs while Case 2 has additional information on the macroeconomic aggregates. Likewise, Table 3a and b demonstrate the results for the CET functions.

The same methodology may also be applied to derive parameters of other parts of the nested CES production structure in the standard CGE model, e.g., between value added and intermediate inputs in the output of each sector or between labor and capital in the value added of each sector. The method is also flexible enough to add more nested structure and can be applied to more flexible behavioral specifications such as the translog functions. In any case, the CES formulation may be viewed as local approximation of a more flexible form (Perroni and Rutherford 1995).

Table 2a: Armington elasticity, using SAMs with limited additional price data, South Korea (1990-2011)

Sectors	Priors or Assumption	Case 1: SAMS deflated by single GDP (3 error support elements)	Case 2: SAMS deflated by GDP deflator and price indices for macro aggregates (3 error support elements)
Agriculture	3.5	3.28	3.39
Mining	3.5	6.17*	6.43*
Manufacturing	3.5	5.55*	5.63*
Utilities	3.5	3.46	3.46
Construction	3.5	3.50	3.50
Services	3.5	0.35*	0.35*

Source: Authors' calculations.

*Posterior estimate is significantly (at 95% of confidence) different from the prior based on a χ^2 goodness or fit test.

Table 2b: Armington elasticity, using SAMs with limited additional price data, SSA (2004-2011)

Sectors	Priors or Assumption	Case 1: SAMS deflated by single GDP (5 error support elements)	Case 2: SAMS deflated by GDP deflator and price indices for macro aggregates (5 error support elements)
Agriculture	2.5	1.14*	0.76*
Mining	2.5	0.25*	0.25*
Manufacturing	3.5	3.98	4.13
Utilities	2.5	2.46	2.35
Construction	2.5	2.52	2.53
Services	2.5	2.03	1.95

Source: Authors' calculations.

*Posterior estimate is significantly (at 95% of confidence) different from the prior based on a χ^2 goodness or fit test.

Table 3a: CET elasticity, using SAMs with limited additional price data, South Korea (1990-2011)

Sectors	Prior estimate or assumption	Case 1: SAMS deflated by single GDP (3 error support elements)	Case 2: SAMS deflated by GDP deflator and price indices for macro aggregates (3 error support elements)
Agriculture	3.5	3.34	3.37
Mining	3.5	3.48	3.48
Manufacturing	3.5	0.35*	0.35*
Utilities	3.5	4.92	5.08
Construction	3.5	3.51	3.50

Services	3.5	0.35*	0.35*
----------	-----	-------	-------

Source: Authors' calculations.

*Posterior estimate is significantly different from the prior based on the $\hat{\chi}^2$ or classical goodness or fit test by Pearson.

Table 3b: CET elasticity, using SAMs with limited additional price data, SSA (2004-2011)

Sectors	Prior estimate or assumption	Case 1: SAMS deflated by single GDP (5 error support elements)	Case 2: SAMS deflated by GDP deflator and price indices for macro aggregates (5 error support elements)
Agriculture	2.5	2.30	1.99
Mining	2.5	0.25*	0.25*
Manufacturing	3.5	3.37	3.41
Utilities	2.5	2.38	1.72
Construction	2.5	2.52	2.52
Services	2.5	2.84	2.79

Source: Authors' calculations.

* Posterior estimate is significantly different from the prior based on the $\hat{\chi}^2$ or classical goodness or fit test by Pearson.

Structural change and productivity

CGE models are increasingly applied to economic scenarios over a long-term horizon. With more SAMs, the approach provides the information basis to calibrate anticipate the economic transformation over the long-term in conjunction or consistent with the elasticity estimation. By economic transformation, we include both shifts in the economic structure and changes in productivity.

A growth rate r may be estimated for the shift or scale parameter in the CES function:

$$A_t = A_{t_0} (1 + r_t)^{t-t_0}$$

The growth in A^t affects both CES factor inputs, but one of the CES factor input may have a factor-specific productivity change, which can also be estimated from our cross entropy methodology:

$$\lambda_t = \lambda_{t_0} (1 + g_t)^{t-t_0}$$

Recent literature seems to suggest that technical change in advanced countries appears to be net factor or labor augmenting (e.g. Jorgenson 2001, Krusell et al. 1997, Carraro and Cian 2009, etc.). The factor augmenting productivity change may be estimated for various CES nested level of the supply side. Table 4 shows the results of our estimates of the growth in TFP and the labor augmenting productivity in the value added function.

In our approach, it is also possible to consider that economic transformation will bring about less rigidity in the production and demand behavior of a developing country. The assumption of a constant elasticity in the Armington and CET functions may not hold over long-run simulations and can be relaxed. One way to handle this is to allow the elasticities to rise over time. If data are insufficient to estimate a flexible functional form, it may still be possible to

incorporate changes in the elasticities if there are sufficient SAMs to cover several growth episodes.

$$\sigma_t = \sigma_{t_0}(1 + h_t)^{t-t_0}$$

$$\Omega_t = \Omega_{t_0}(1 + i_t)^{t-t_0}$$

Alternatively, foreign trade shares can be allowed to change to capture the effects of globalization, effects that are difficult to capture in a homothetic function like the CES or CET specification. A different approach is needed to capture rising trade shares as output expands. One option is to calibrate the change in the CES share or distribution parameter. In our approach, a share parameter corresponding to each year of SAM can be computed simultaneously with the other parameters and its growth rate can be derived:

$$\delta_t = \delta_{t_0}(1 + \phi_t)^{t-t_0}$$

Importantly, there will be constraints in the number of parameters to be estimated simultaneously in our system approach due to the degrees of freedom reflected in the behavioral relationships and in the limited amount of data. The estimates are therefore linked intrinsically. For example, the posterior estimates of the elasticities will tend to match their prior values if many of the other parameters are set free and/or are capturing much of the structural change story in the economy. The trade-offs can be exploited in the following ways. If there are good prior estimates of the elasticities, then the entropy method can be used to gain better estimates of the productivity and other parameters. If new estimates of the elasticities are important, then no many of the others can be set free.

To illustrate, the estimates of the Armington and CET elasticities in Table 2 to 3 are consistent with what are being allowed to be estimated with respect to the other parameters in the CGE model. In both cases, we also allow for the estimation of the labor augmenting productivity. In the first year, one sector, services in South Korea and mining in SSA, is set as 1.0 (reference), allowing other sectors to vary relative to it so that there are relative differences in the sector productivity to start with. In subsequent years, labor augmenting productivity in all sectors will vary relative to their initial values. In addition, the TFP, the scale parameters (shifters) in the other CES type functions and the respective share parameters are kept the same across time at their respective levels in the first year.

Table 4a: Total factor productivity and labor augmenting productivity, South Korea

Sectors	TFP annual compound growth (in %)	Labor augmenting productivity (LAP) Prior	Labor augmenting productivity (LAP) annual compound growth (in % for CASE 1) (3 error support elements)	Labor augmenting productivity (LAP) annual compound growth (in % for CASE 2) (3 error support elements)
Agriculture	0	1	0.48	-1.05
Mining	0	1	-6.44	-9.14
Manufacturing	0	1	1.42	0.21

Utilities	0	1	-3.69	-1.05
Construction	0	1	-3.27	2.01
Services	0	1	3.36	3.36

Source: Authors' calculations.

Table 4b: Total factor productivity and labor augmenting productivity, SSA

Sectors	Labor augmenting productivity (LAP) Prior value	Labor augmenting productivity (LAP) annual compound growth (in % for CASE 1) (5 error support elements)	Labor augmenting productivity (LAP) annual compound growth (in % for CASE 2) (5 error support elements)
Agriculture	1	-17.00	-17.00
Mining	1	-14.00	-14.00
Manufacturing	1	-13.00	-13.00
Utilities	1	1.26	2.73
Construction	1	-4.99	-4.19
Services	1	11.74	10.19

Source: Authors' calculations.

Table 5a: Elasticity of substitution for various CES functions (Case 2), South Korea) (3 error support elements)

Sectors	Between labor and capital in value added	Between value added and intermediate good in output	Between imports and domestic goods (Armington)	Between exports and domestic goods (CET)
Agriculture	3.14 (2.5)	2.67 (2.5)	3.39 (3.5)	3.37 (3.5)
Mining	2.75 (2.5)	0.34 (2.5)*	6.43 (3.5)*	3.48 (3.5)
Manufacturing	0.84 (2.5)*	0.51 (2.5)*	5.63 (3.5)*	0.35 (3.5)*
Utilities	2.27 (2.5)	1.20 (2.5)	3.46 (3.5)	5.08 (3.5)
Construction	2.51 (2.5)	2.89 (2.5)	3.50 (3.5)	3.50 (3.5)
Services	1.89 (2.5)	0.25 (2.5)*	0.35 (3.5)*	0.35 (3.5)*

Source: Authors' calculations.

Note: Numbers in parenthesis are the priors.

*Posterior estimate is significantly (at 95% of confidence) different from the prior based on a χ^2 goodness or fit test.

Table 5b: Elasticity of substitution for various CES functions (Case 2), SSA) (5 error support elements)

Sectors	Between labor and capital in value added	Between value added and intermediate good in output	Between imports and domestic goods (Armington)	Between exports and domestic goods (CET)
Agriculture	0.62 (1.4)*	0.16 (1.4)*	0.76 (2.5)*	1.99 (2.5)
Mining	1.18 (1.4)	0.20 (1.4)*	0.25 (2.5)*	0.25 (2.5)*
Manufacturing	0.68 (1.4)*	0.16 (1.4)*	4.13 (3.5)	3.41 (3.5)
Utilities	1.32 (1.4)	1.44 (1.4)	2.35 (2.5)	1.72 (2.5)
Construction	0.88 (1.4)	1.41 (1.4)	2.53 (2.5)	2.52 (2.5)
Services	0.34 (1.4)*	1.79 (1.4)	1.95 (2.5)	2.79 (2.5)

Source: Authors' calculations.

Note: Numbers in parenthesis are the priors.

*Posterior estimate is significantly (at 95% of confidence) different from the prior based on a χ^2 goodness or fit test.

Conclusions and Future Research

In this paper, we defined a formal Bayesian approach to estimate parameters of a CGE model under noisy and limited data environment. The cross-entropy estimation method described in the paper is potentially applicable to all calibrated CGE models and their SAM database.

Inferring information about parameters and elements of a CGE model of a developing country is often made difficult because of incomplete, infrequent, or uncertain evidence such as

dated or poorly constructed SAMs and auxiliary data. To estimate behavioral parameters and structural change in CGE models, we implement a two-step cross-entropy estimation method. The data step adjusts the historical SAMs of South Korea and Sub-Saharan Africa to a common base year taking into account that the SAMs are measured with errors and relative price indices are scant. Next, a parameter estimation step calculates (filters) parameters and structural change simultaneously within the specifications of the CGE model postulated the country/region.

The approach can be used to estimate how economic transformation will bring about structural change and less rigidity in the production and demand behavior of a developing country. There are, of course, trade-offs in the number of parameters to be estimated simultaneously due to the degrees of freedom and data constraints. As illustrations, we estimated the Armington and CET elasticities for South Korea and Sub-Saharan Africa that were consistent with the case where labor augmenting productivity captured compositional changes in the sector value-added and indirectly, its effects on the final demand vectors. There are other possibilities, such as allowing for trade elasticities and trade shares to evolve in order to capture the effects of globalization and increasing trade.

We have a few suggestions for future research. One is to do a systematic analysis of different combination of the critical parameters and examine how well they improve the estimation of trade elasticities and other factors of structural change. A second research is to employ extraneous estimates of historical productivity change and trade pattern in order to improve the estimates of trade elasticities and/or their evolution. Lastly, the approach seems ideal for regional or global CGE modeling where SAM and other auxiliary data are still very limited; hence, we hope to extend the estimation of parameters to various regions outside of Sub-Saharan Africa.

I. References

- ARNDT, C., S. ROBINSON, and F. TARP (2002): "Parameter Estimation for a Computable General Equilibrium Model: A Maximum Entropy Approach," *Economic Modelling*, 19, 375-398.
- ARROW, K. J., H. B. CHENERY, B. S. MINHAS, and R. M. SOLOW (1961): "Capital-Labor Substitution and Economic Efficiency," *The Review of Economics and Statistics*, 225-250.
- CHENERY, H. B., and H. ELKINGTON (1979): *Structural Change and Development Policy*. Washington D.C.: The World Bank and Oxford University Press.
- CHENERY, H. B., S. ROBINSON, and M. SYRQUIN (1986): *Industrialization and Growth*. New York: Oxford University Press.
- CHENERY, H. B., M. SYRQUIN, and H. ELKINGTON (1975): *Patterns of Development, 1950-1970*. New York: The World Bank and Oxford University Press.

- CICOWIEZ, M., and H. LOFGREN (2006): *Cge Model: Single Country with Open Economy* CEDLAS - Universidad Nacional de la Plata, Buenos Aires, Argentina (unpublished manuscript).
- DAWKINS, C., T. N. SRINIVASAN, and J. WHALLEY (2001): "Calibration," *Handbook of econometrics*, 5, 3653-3703.
- DE MELO, J., and S. ROBINSON (1989): "Product Differentiation and the Treatment of Foreign Trade in Computable General Equilibrium Models of Small Economies," *Journal of International Economics*, 27, 47-67.
- DERVIS, K., J. DE MELO, and S. ROBINSON (1982): *General Equilibrium Models for Development Policy*. Cambridge, UK: Cambridge University Press.
- GOLAN, A., G. JUDGE, and D. MILLER (1996): *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. John Wiley.
- JORGENSEN, D. W. (2011): "Innovation and Productivity Growth," *American Journal of Agricultural Economics*, 93, 276-296.
- JORGENSEN, D. W., M. S. HO, and J. D. SAMUELS (2012): "A Prototype Industry-Level Production Account for the United States, 1947-2010."
- JORGENSEN, D. W., and M. P. TIMMER (2011): "Structural Change in Advanced Nations: A New Set of Stylised Facts*," *The Scandinavian Journal of Economics*, 113, 1-29.
- JORGENSEN, D. W., and K.-Y. YUN (2013): "Taxation, Efficiency and Economic Growth," in *Handbook of Computable General Equilibrium Modeling*, ed. by P. B. Dixon, and D. W. Jorgenson. MA, USA: North-Holland, 659-740:.
- JUDGE, G. G., and R. C. MITTELHAMMER (2012): *An Information Theoretic Approach to Econometrics*. Cambridge, UK: Cambridge University Press.
- LÖFGREN, H., R. L. HARRIS, and S. ROBINSON (2002): *A Standard Computable General Equilibrium (Cge) Model in Gams*. Washington, D.C.: International Food Policy Research Institute.
- MCGRAYNE, S. B. (2011): *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines & Emerged Triumphant from Two Centuries of Controversy*. New Heaven, Connecticut: Yale University Press.
- Okushima, S. and M. Tamura (2009). "A Double Calibration Approach to the Estimation of Technological Change." *Journal of Policy Modeling* 31(2009): 119-125.
- PERRONI, C., and T. F. RUTHERFORD (1995): "Regular Flexibility of Nested Ces Functions," *European Economic Review*, 39, 335-343.
- PYATT, G., and J. I. ROUND (1985): "Social Accounting Matrices: A Basis for Planning."

- ROBINSON, S., A. CATTANEO, and M. EL-SAID (2001): "Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods," *Economic Systems Research*, 13, 47-64.
- STONE, R. (1962): *A Social Accounting Matrix for 1960*. Department of Applied Economics. University of Cambridge.
- Zellner, A. (1988). "Optimal Information Processing and Bayes's Theorem." *The American Statistician*. Vol 42, Issue 4, pp. 278-280.

Appendix 1: figures and tables

Figure 1. CGE structure: production & consumption

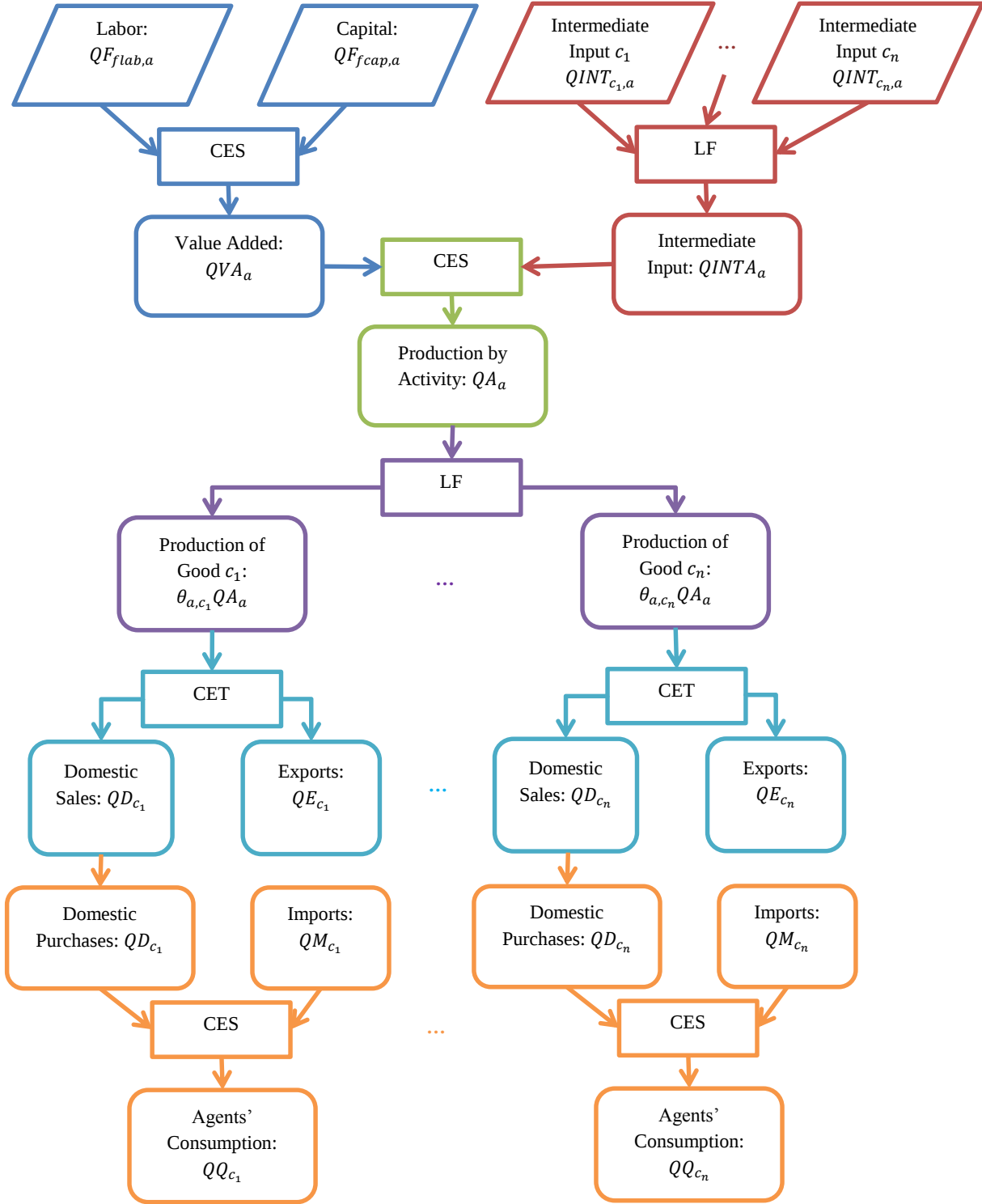


Figure 2. Korean factor elasticity – labor productivity – shares graphs

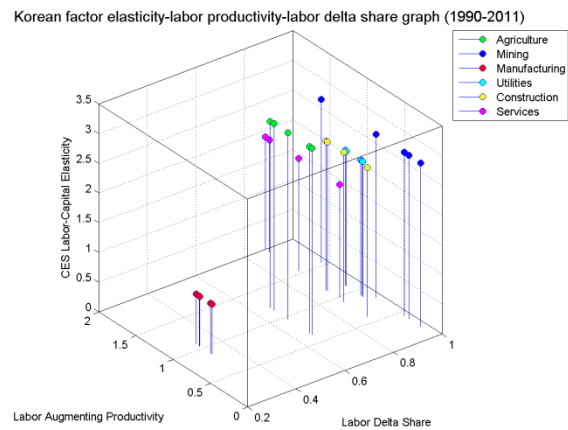


Figure 3. SSA factor elasticity – labor productivity – delta shares graphs

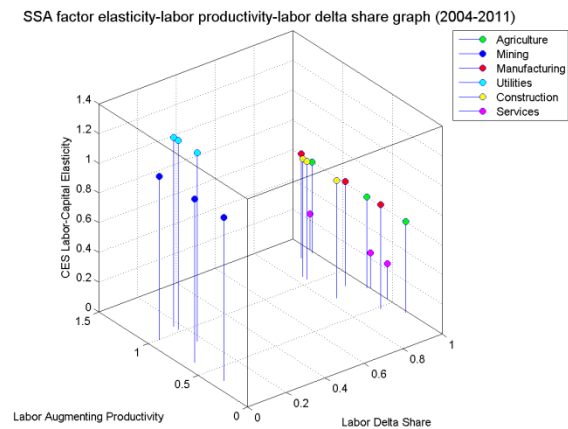


Figure 4. Korean factor productivity – factor income share graphs

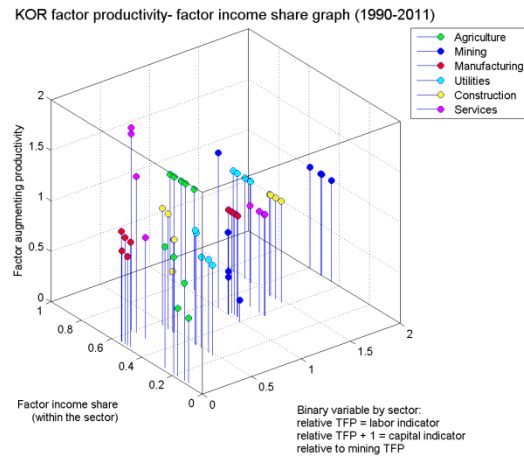
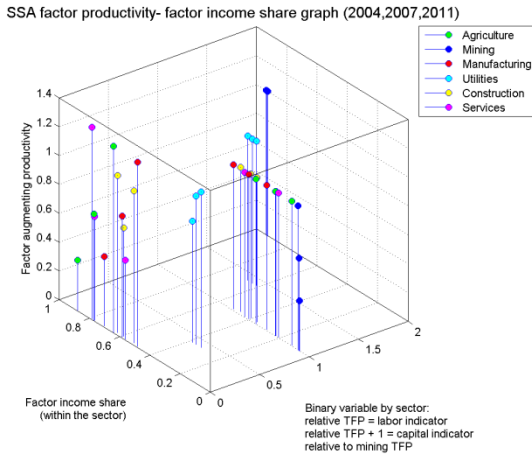


Figure 5. SSA factor productivity – factor income share graphs



Appendix 2: Some Symbols and Notations

Indices

k_m	an index for the outcomes (support) of the m^{th} behavioral parameter; $k_m \in \{2, 3, \dots, K: K \in \mathbb{N}_{>1}\} \forall m$
k_n	an index that embodies the outcomes (support) of the n^{th} SAM target; $k_n \in \{2, 3, \dots, K: K \in \mathbb{N}_{>1}\} \forall n$
k_s	an index for the outcomes (support) of the s^{th} unobserved parameter; $k_s \in \{2, 3, \dots, K: K \in \mathbb{N}_{>1}\} \forall s$
m	an index for a behavioral parameter on the $\mathbf{B}$ matrix; $m \in \{0, 1, 2, \dots, M: M \in \mathbb{N}_0\}$
n	a subindex for a SAM target on the $\mathbf{Y}$ matrix; $n \in \{0, 1, 2, \dots, N: N \in \mathbb{N}_0\}$
n_t^{SAM}	the number of selected SAM targets at period t
s	an index of an unobserved (technical change) parameter on the $\mathbf{Z}$ matrix; $s \in \{0, 1, 2, \dots, S: S \in \mathbb{N}_0\}$
t	a time (period) index; $t \in \{t_1, \dots, t_T: T \in \mathbb{N}_{>1}\}$

Vectors and matrices

δ	a vector of calibration parameters, which is a second vector of behavioral parameters whose values are uniquely determined by the selection of the $\mathbf{B}$ values, the exact form of $\mathbf{F}(\cdot)$, and the data in the base year t_0
Φ	a vector valued function producing the calibration δ parameters
$\mathbf{B}$	an M -dimensional vector of behavioral parameters such as the Armington elasticities
$\mathbf{F}$	an $(I * T)$ -dimensional vector valued function
$\mathbf{G}$	a $(L * T)$ -dimensional vector valued function
$\mathbf{V}$	a $(L * L * T)$ -matrix representation of the SAMs
$\mathbf{X}_t$	an I -dimensional vector of endogenous CGE variables such as prices and quantities
$\mathbf{Y}$	a $(n_{t_1}^{SAM} + n_{t_2}^{SAM} + \dots + n_{t_T}^{SAM})$ -dimensional vector of SAM targets
$\mathbf{Z}$	an $(S + R)$ -dimensional vector of exogenous parameters such as endowments and tariff rates; $\mathbf{Z}$ is partitioned in two components: $\mathbf{Z} = \{\mathbf{Z}_t^u, \mathbf{Z}_t^o\}$
$\mathbf{Z}_t^o$	an R -dimensional vector of observed exogenous parameters, which may consist of historical data elements such as tax rates, endowments, prices, government spending, household consumption or saving rates, etc.
$\mathbf{Z}_t^u$	an S -dimensional vector of unobserved parameters, which may contain rates of technical change, implicit or unknown tax or subsidy rates, and other items that are not available from the historical information

Errors and probabilities

$e_{B,m}$	the posterior error value for the m^{th} behavioral parameter; $e_{B,m} \in \mathbb{R} \forall m$
-----------	---

$e_{Y,t,n}$ the posterior error value for the n^{th} SAM target flow at period t ; $e_{Y,t,n} \in \mathbb{R} \ \forall n, \ \forall t$
 $e_{Z,t,s}$ the posterior error value for the s^{th} unobserved parameter at time t ; $e_{Z,t,s} \in \mathbb{R} \ \forall s, \ \forall t$
 v_{B,k_m} the k^{th} error outcome for the m^{th} behavioral parameter; $v_{B,k_m} \in \mathbb{R} \ \forall k_m$
 v_{Y,t,k_n} the k^{th} error outcome for the n^{th} SAM target at period t ; $v_{Y,t,k_n} \in \mathbb{R} \ \forall k_n, \ \forall t$
 v_{Z,t,k_s} the representation of the k^{th} error outcome for the s^{th} unobserved parameter at period t ;
 $v_{Z,t,k_s} \in \mathbb{R} \ \forall k_s, \ \forall t$
 w_{B,k_m} the posterior probability (weight) of error outcome v_{B,k_m} ; $w_{B,k_m} \in [0,1] \ \forall k_m$
 w_{Y,t,k_n} the posterior weight of error outcome v_{Y,t,k_n} ; $w_{Y,t,k_n} \in [0,1] \ \forall k_n, \ \forall t$
 w_{Z,t,k_s} the posterior probability (weight) of error outcome v_{Z,t,k_s} ; $w_{Z,t,k_s} \in [0,1] \ \forall k_s, \ \forall t$
 $\bar{w}_{B,k_m}$ the prior probability (weight) of error outcome v_{B,k_m} ; $\bar{w}_{B,k_m} \in [0,1] \ \forall k_m$
 $\bar{w}_{Y,t,k_n}$ the prior probability (weight) of error outcome v_{Y,t,k_n} ; $\bar{w}_{Y,t,k_n} \in [0,1] \ \forall k_n, \ \forall t$
 $\bar{w}_{Z,t,k_s}$ the prior probability (weight) of the error outcome v_{Z,t,k_s} ; $\bar{w}_{Z,t,k_s} \in [0,1] \ \forall k_s, \ \forall t$

Appendix 3: Prior Error Distributions

Error Specification

The typical specification of the error is either multiplicative or additive:

$$X = \bar{X} + e \text{ or } \ln(X) = \ln(\bar{X}) + e$$

where $\bar{X}$ is the prior mean and e is the error term with mean zero.

The error is written as the discrete probability weighted sum of an error support set.

$$e = \sum_k W_k \cdot \bar{V}_k$$
$$0 \leq W_k \leq 1 \text{ and } \sum_k W_k = 1$$

The V parameters are fixed and have the units of the item (X) being estimated. They define the domain of possible values that X can take and so contain information for estimation. The W s are discrete probabilities, defining the probability mass function for the distribution of the error. This specification converts the problem of estimating errors in “natural” units into a Bayesian problem of estimating a set of probabilities. Instead of directly estimating the mean and variance of a random variable, we are now estimating a discrete probability distribution. The specified support set provides the link.

The number of elements in the support set, k , defines the number of probability weights (W) that need to be estimated and will determine the parameters of the distribution of the errors that can be “recovered” from the data. To estimate the errors, one starts with the error support set (V) and a prior on the probability weights ($\bar{W}$). This prior can be “uninformative” or “informative”, depending on the choice of the prior probability weights.

Uninformative Prior

An uninformative prior provides information only about the bounds between which the errors must be located. In Bayesian estimation and information theory, the most uninformative prior is the uniform distribution, which has maximum entropy. We will specify a discrete prior probability mass function that approximates the continuous uniform distribution between known upper and lower bounds.

Assume that the upper and lower bounds on V are given by plus or minus $3s$, where “ s ” is a specified constant. For the continuous uniform distribution between these bounds, the variance is:

$$\sigma^2 = \frac{(3s - (-3s))^2}{12} = 3s^2$$

Specify an evenly-spaced, 7-element support set ($k = 7$) with identical (uniform) prior probability weights:

$$\begin{aligned}\bar{V}_1 &= -3s & \bar{V}_2 &= -2s & \bar{V}_3 &= -s & \bar{V}_4 &= 0 \\ \bar{V}_5 &= +s & \bar{V}_6 &= +2s & \bar{V}_7 &= +3s \\ \sigma^2 &= \sum_k \bar{W}_k \cdot \bar{V}_k^2 & \text{and the prior is } \bar{W}_k &= \frac{1}{7} \\ \sigma^2 &= \frac{s^2}{7} (9 + 4 + 1 + 1 + 4 + 9) = 4s^2\end{aligned}$$

A discrete uniform prior with 7-element support set is a conservative uninformative prior, with a prior variance of $4s^2$. Adding more elements to the support set would more closely approximate the continuous uniform distribution, reducing the prior variance toward the limit of $3s^2$. Note that the estimated posterior distribution will be essentially unconstrained.

Informative 2-Parameter Prior

Start with a prior on both the mean and standard deviation of a symmetric, two-parameter error distribution. The prior mean on the error is zero by construction and the prior standard deviation of e is specified as the prior on the standard error of measurement of the item. Specify an evenly-spaced support set with $s=\sigma$ so that the bounds are now $\pm 3\sigma$.

$$\begin{aligned}\sum_k \bar{W}_k \cdot \bar{V}_k &= 0 & \text{mean} \\ \sum_k \bar{W}_k \cdot \bar{V}_k^2 &= \sigma^2 & \text{variance} \\ \bar{V}_1 &= -3\sigma, \bar{V}_2 = 0 \text{ and } \bar{V}_3 = +3\sigma\end{aligned}$$

$$\begin{aligned}0 &= \bar{W}_1 \cdot (-3\sigma) + \bar{W}_2 \cdot (0) + \bar{W}_3 \cdot (3\sigma) & \text{mean} \\ \sigma^2 &= \bar{W}_1 \cdot (+9\sigma^2) + \bar{W}_2 \cdot (0) + \bar{W}_3 \cdot (9\sigma^2) & \text{variance} \\ \bar{W}_1 &= \bar{W}_3 = \frac{1}{18} & \text{symmetry} \\ \bar{W}_2 &= 1 - \bar{W}_1 - \bar{W}_3 = \frac{16}{18}\end{aligned}$$

Estimation of a posterior distribution in this case can retrieve information about essentially two moments of the error distribution, since the 3-element prior only allows two degrees of freedom in estimation (since the probability weights must sum to one). One can specify a more informative prior using a larger support set.

Informative 4-Parameter Prior

To recover more information about the error distribution the prior must include more moments—for example: mean, variance, skewness, and kurtosis. Assume a normal distribution with a prior for the mean and variance so that prior skewness is zero and kurtosis is a function of σ .

$$\begin{aligned}\sum_k \bar{W}_k \cdot \bar{V}_k &= 0 && \text{mean} \\ \sum_k \bar{W}_k \cdot \bar{V}_k^2 &= \sigma^2 && \text{variance} \\ \sum_k \bar{W}_k \cdot \bar{V}_k^3 &= 0 && \text{skewness} \\ \sum_k \bar{W}_k \cdot \bar{V}_k^4 &= 3\sigma^4 && \text{kurtosis}\end{aligned}$$

Specify an evenly-spaced 5-element support set:

$$\bar{V}_1 = -3.0\sigma, \bar{V}_2 = -1.5\sigma, \bar{V}_3 = 0, \bar{V}_4 = +1.5\sigma, \bar{V}_5 = +3.0\sigma,$$

In this case, the prior on the probability weights can be calculated from the known prior moments and the assumption of symmetry:

$$\begin{aligned}0 &= \bar{W}_1 \cdot (-3) + \bar{W}_2 \cdot (-1.5) + \bar{W}_3 \cdot (0) + \bar{W}_4 \cdot (1.5) + \bar{W}_5 \cdot (3) && \text{(mean)} \\ \sigma^2 &= \bar{W}_1 \cdot (9\sigma^2) + \bar{W}_2 \cdot \left(\frac{9}{4}\sigma^2\right) + \bar{W}_3 \cdot (0) + \bar{W}_4 \cdot \left(\frac{9}{4}\sigma^2\right) + \bar{W}_5 \cdot (9\sigma^2) && \text{(variance)} \\ 0 &= \bar{W}_1 \cdot (-27\sigma^3) + \bar{W}_2 \cdot \left(-\frac{27}{8}\sigma^3\right) + \bar{W}_3 \cdot (0) + \bar{W}_4 \cdot \left(\frac{27}{8}\sigma^3\right) + \bar{W}_5 \cdot (27\sigma^3) && \text{(skewness)} \\ 3\sigma^4 &= \bar{W}_1 \cdot (81\sigma^4) + \bar{W}_2 \cdot \left(\frac{81}{16}\sigma^4\right) + \bar{W}_3 \cdot (0) + \bar{W}_4 \cdot \left(\frac{81}{16}\sigma^4\right) + \bar{W}_5 \cdot (81\sigma^4) && \text{(kurtosis)} \\ \bar{W}_1 = \bar{W}_5 &= \frac{1}{162}; \quad \bar{W}_2 = \bar{W}_4 = \frac{16}{81}; \quad \bar{W}_3 = \frac{48}{81}\end{aligned}$$

As with the other priors, the estimated posterior distribution is unconstrained.