ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT

Journal of Econometrics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

JOURNAL OF Econometrics

www.elsevier.com/locate/jeconom

Indirect inference and calibration of dynamic stochastic general equilibrium models

Ramdan Dridi, Alain Guay[a,*], Eric Renault[b,c]

[a]Dept. des Sciences Economiques, Université du Québec à Montréal, CIRPÉE and CIREQ, Case postale 8888, Succursale Centre-Ville, Montreal, Qué., Canada H3C3P8

[b]University of North Carolina, USA

[c]CIRANO and CIREQ, Canada

Abstract

We advocate in this paper the use of a sequential partial indirect inference (SPII) approach, in order to account for calibration practice where dynamic stochastic general equilibrium models (DGSE) are studied only through their ability to reproduce some well-chosen moments. We stress that, despite a lack of statistical formalization, the controversial calibration methodology addresses a genuine issue on the consequences of misspecification in highly nonlinear and dynamic structural macro-models. We argue that a well-driven SPII strategy might be seen as a rigorous calibrationnist approach, that captures both the advantages of this approach (accounting for structural "a-statistical" ideas) and of the inferential approach (precise appraisal of loss functions and conditions of validity). This methodology should be useful for the empirical assessment of structural models such as those stemming from the real business cycle theory or the asset pricing literature.

JEL classification: C52; C51; C15

Keywords: Calibration; Indirect inference; Structural models; Real business cycle; Asset pricing

1. Introduction

Dynamic stochastic general equilibrium (DSGE) models are the common framework of new classical macroeconomics, with the ambition to provide structural microfoundations

*Corresponding author.

E-mail addresses: guay.alain@uqam.ca (A. Guay), renault@email.unc.edu (E. Renault).

0304-4076/$ - see front matter © 2005 Elsevier B.V. All rights reserved.

doi:10.1016/j.jeconom.2005.11.003

for macroeconomics. However, this ambition comes at a price. Nobody can believe that DSGE models present a descriptively realistic model of the economic process. "Of course, the model is not 'true'" (Lucas, 1987) and this is probably the reason why the advent of DSGE models has led new classical macroeconomics to turn to calibration methods as an alternative to classical econometrics, involving estimation and testing.

The endorsement of calibration as an alternative to estimation, and the related endorsement of verification as an alternative to statistical tests may lead to the conclusion that "the new classical macroeconomics is now divided between calibrators and estimators" (Hoover, 1995). However some econometricians claim that considering, as Lucas (1987) and Kydland and Prescott (1991) do that "the specification errors being committed are of sufficient magnitude as to make conventional estimation and testing of dubious value" is simply misunderstanding econometrics since "traditional model building never proceeded under the assumption that any model was true" (Kim and Pagan, 1995).

The approach we advocate in this paper is somewhere in between the two extreme views that either some unrealistic features of DSGE models should lead to eschew orthodox econometrics altogether or that calibrators simply misunderstand that traditional econometrics "never proceed under the assumption that any model was true". On the contrary, we do think that econometricians have something to learn from calibrators and we try to go further in the research program put forward by Hansen and Heckman (1996): "model calibration and verification can be fruitfully posed as econometric estimation and testing problems".

We argue, by contrast with the "never" claim above, that more often than not econometric practices are seriously flawed with a maintained assumption of model truth. The recent regain of popularity of maximum likelihood (MLE) approaches to DSGE precisely shows that many econometricians still consider that MLE is the best thing to do, at least when it is tractable. However, there is no such thing in econometric theory as compelling arguments in favor of MLE in the case of misspecified models. Of course, properties of MLE in case of misspecification, also called quasi- or pseudo-maximum likelihood (QMLE) are well known since White (1982) and Gouriéroux et al. (1984). However, while the former stresses that QMLE converges towards a pseudo-true value of the unknown parameters and that its asymptotic variance is no longer conformable to the common Cramer Rao bound but must be replaced by the so-called sandwich formula, the latter characterizes the very restrictive assumptions under which the pseudo-true value coincides with the true unknown value. In other words, not only QMLE does not provide such thing as an efficient asymptotic variance but, even worse, it leads to select a pseudo-true value of unknown parameters which may be quite different from the one which would be associated to an economically meaningful loss function.

The econometrician's hopeless search for a well-specified parametric model ("quest for the Holy Grail" as dubbed by Monfort (1996)) and associated efficient estimators even remain popular when MLE becomes intractable due to highly nonlinear dynamic structures including latent variables. Efficiency properties of "efficient method of moments" (EMM, Gallant and Tauchen, 1996) or more generally of generalized method of moments (GMM, Hansen, 1982), simulated method of moments (SMM, Duffie and Singleton, 1993) and indirect inference (II, Gouriéroux et al. (1993)) when the set of moment conditions is sufficiently large to span the likelihood scores are often advocated as if the likelihood score was something well specified. Actually, not only one should not forget that we are the most often dealing with a pseudo-score but the resort to simulation requires even more care since the likely misspecified structural parametric model is used as a simulator.

This paper is a contribution to the econometric literature that has "attempted to tame calibration and return it to the traditional econometric fold" by interpreting "calibration as a form of estimation by simulation" (Hoover, 1995) along the lines of Manuelli and Sargent (1988), Gregory and Smith (1990), Canova (1994) and Bansal et al. (1995). However, even more focus is put on the likely severe misspecification of structural models stemming from the DSGE literature. This leads us to an explicit account of calibrators' recommendations, while showing that they may be made compatible with a well-established approach to econometrics. In other words, we aim at delineating a close methodology which could be able to gather both the advantages of the inferential approach (estimation, confidence sets and specification testing) and also the advantages of the calibration approach that correspond, in our opinion, to consistent estimation of some structural parameters of interest and robust prediction and induction despite misspecification of the structural model.

Contributions of this paper are threefold.

First, we point out that asymptotic variance formulas for any kind of simulated moment-based method (SMM, EMM or II) must take into account some kind of sandwich formulas for the choice of efficient weighting matrices and associated formulas for asymptotic variance of estimators. Forgetting this kind of correction is even more detrimental than for QMLE since two kinds of sandwich formulas must be taken into account, one for the data generating process (DGP) and one for the simulator which turns out to be different from the DGP in case of misspecification. Moreover, since only endogenous variables are simulated, correct formulas for asymptotic variance matrices require a specific account for exogenous variables. In this respect, we extend the results of Gouriéroux et al. (1993) theory of II to a case of possible misspecification of the simulator.

As for QMLE, misspecification may not only imply a violation of standard asymptotic variance formulas but even more importantly, may lead the econometrician to consistently estimate a pseudo-true value which may have nothing to do with the true unknown value of the parameters of interest. The second contribution of this paper is to put forward the encompassing tests methodology as a way to focus SMM or more generally II estimators on the consistent estimation of the true unknown value θ_1^0 of a subset θ_1 of the full set $\theta = (\theta_1, \theta_2)$ of structural parameters. While a fully parametric model, that is a family of probability distributions indexed by $\theta = (\theta_1, \theta_2)$ is needed to get a simulator, there is no hope to find any economic theoretical underpinnings for such parametric DSGE models which cannot be more than a crude idealization of the economic process. Unfortunately, the matching moment strategy of estimation is an indirect approach to inference about the structural parameters θ which goes through a binding function $\beta(\theta)$ relating the structural parameters θ to some instrumental parameters β which can be directly estimated from their sample counterparts. Note that in this respect, II approach to nonlinear analytically intractable structural models is nothing but an extension of the old indirect least-squares approach to linear simultaneous equations models. In our nonlinear and misspecified structural model context, it is unfortunately highly hazardous to get a consistent estimator of the true unknown value θ_1^0 of a subset θ_1 when solving with respect to θ_1 a sample and possibly simulation-based counterpart of the equations $\beta(\theta_1, \theta_2^*) = \beta^0$ where β^0 denotes the true unknown value of the instrumental parameters β (by definition easy to estimate) but θ_2^* is only a pseudo-true value of θ_2. The necessary condition, that is $\beta^0 = \beta(\theta_1^0, \theta_2^*)$

precisely means that the structural model, albeit misspecified, encompasses the instrumental one.

The requirement of encompassing typically means that, if we do not want to proceed under the maintained assumption that the structural model is true, we must be parsimonious with respect to the number of moments to match or more generally to the scope of macroeconomic evidence that is captured by the instrumental model, as parameterized by β, like for instance the coefficients of a vector autoregression. This is at odds with the efficiency kind of goal as advocated by Bansal et al. (1995) to endorse the EMM approach to calibration: "if a structural model is to be implemented and evaluated on statistical criteria i.e. one wants to take seriously statistical test and inference, the structural model has to face all empirically relevant aspects of the data". We are not far to think on the contrary like Prescott (1983) that "if any observation can be rationalized with some approach, then that approach is no scientific" or at least like Lucas (1980) that "insistence on the 'realism' of an economic model subverts its potential usefulness in thinking about reality". Economic reality may be interestingly captured by the parameters of interest θ_1 while there is no hope to find the Holy Grail of a fully parametric true model indexed by (θ_1, θ_2). Then, as often stressed by calibrators, it is important to have in mind a hierarchy of moments, with first place given to some specific βs like unconditional means, variances and correlations rather than more sophisticated characteristics of conditional probability distributions. The key point is that while a true parametric model defining a true unknown value (θ_1^0, θ_2^0) would by definition ensure the necessary encompassing condition, whatever the dimension of β (even with at the limit an infinite dimensional vector β of auxiliary parameters as for EMM), the equations $\beta(\theta_1, \theta_2^*) = \beta^0$ are going to characterize the true unknown θ_1^0 whatever the misspecification about θ_2, only if we have chosen a convenient instrumental model which does not capture what goes wrong in the paths simulated from the structural model endowed with the fictitious value (θ_1, θ_2^*) of the structural parameters. This is the reason why we advocate in this paper the partial indirect inference (PII) approach.

PII is well suited in case of partial encompassing. It means that only a subset of the encompassing equations $\beta(\theta_1^0, \theta_2^*) = \beta^0$ appear to be fulfilled. By restricting ourselves to such a subset, we may have to renounce to the complete identification of the vector θ of structural parameters. By contrast to a narrow view of econometric identification, this is typically something we can accept insofar as underidentification is only about some "pseudo-parameters" θ_2, that is to say quantities which are known to be poorly related to economic reality, as captured by our structural model. Then, as calibrators do, we propose to fix the value of these unidentified parameters to some "reasonable" levels. These "calibrated" values are needed to perform simulations for the determination of the binding function but do not contaminate a subset of equations for which the encompassing property turns out to be fulfilled. In other words, we find a rationale to the calibration practice within a well founded econometric methodology. A good reason not to apply a neutral moment matching to identify all the parameters is that it is along only some selected dimensions that we may hope to get meaningful quantitative assessments from our structural model. For example, as reminded by Hansen and Heckman (1996), some "particular time series frequencies could be deemphasized in adopting an estimation criterion because misspecification of a model is likely to contaminate some frequencies more than others (Hansen and Sargent, 1993)". By still seeking econometric identification of all structural parameters θ_1 and θ_2, the econometrician runs the risk to contaminate the estimation of the parameters of interest θ_1 with the likely misspecification of the part of the model concerning θ_2. Amazingly, our PII kind of extension of Gouriéroux et al. (1993) theory of II fully concurs, even in the terminology, with the Hoover characterization of Lucas (1980) and Prescott (1983) ‘‘discipline of the calibration method’’: it ‘‘comes from the paucity of free parameters (...) in some sense, the calibration method would appear to be a kind of indirect estimation’’. We claim more precisely that it is because the estimation of structural models is generally ‘‘indirect’’, in the sense that it takes a binding function relating structural parameters to instrumental ones, that calibration matters to pin down some ‘‘key parameters’’ θ_2 from calibrator’s knowledge rather than from an orthodox moment matching procedure. These key parameters are so because they define some components of θ_2, which prevent us from getting full encompassing an thus to estimate consistently the parameters of interest θ_1, when contaminated by the identification of θ_2.

A third contribution of this paper is to propose a sequential approach to PII, in order to accommodate not only the calibration step but also the verification step of the common empirical practice for DSGE. More precisely, we do think as calibrators that the specification tests should only be focused on the reproduction of stylized facts the structural model is aimed to reproduce. But our additional discipline amounts to a second step of specification testing, once the parameters of interest θ_1 have been hopefully consistently estimated in a first step from matching moment simulated with a possibly calibrated θ_2. The second step of simulated moment matching (or minimization of any kind of economically meaningful loss function) with respect to these previously calibrated components aims at controlling the degree of misspecification at a reasonable level, that is there is no such thing like a gross inability of our structural model to reproduce the economically meaningful moments. Since the procedure is a two step one, we call it a sequential partial indirect inference (SPII). In our opinion, this two step simulated moments matching methodology remains exactly true to the calibrators’ point of view: reproducing some dimensions of interest under the constraint that some structural parameters of interest are consistently estimated. This is precisely because the requirement of consistency is maintained that the two steps are disentangled whatever the cost in terms of efficiency of a two-step procedure of estimation. Of course, if the structural model were well specified, a one step estimator of θ_1 and θ_2 jointly would be preferable. The aim of roughly reproducing broad economic reality of interest must not make us running the risk of inconsistently estimating the crucial structural parameters. Otherwise, it would be a purely data-based approach. We claim on the contrary (see e.g. our reinterpretation below of the Mehra and Prescott (1985) equity premium puzzle exercise) that consistent estimation of a few structural parameters is a binding constraint for calibrators. The second step of verification, as we perform it, is consistent with the Canova (1994) kind of interpretation of the calibration practice. The question asked is: ‘‘Given that the model is false, how true is it?’’.

As already mentioned, this paper is far to be the first to address the issue of a statistical appraisal of the calibration methodology. However, only a few papers have focused on the consequences of misspecification in simulated moments matching. While intriguing Bayesian approaches to calibration of misspecified models have been proposed by Canova (1994), Dejong et al. (1996), Geweke (1999) and Schorfheide (2000), we argue that SPII is the convenient way to accommodate it with a frequentist point of view.

The paper is organized as follows. In Section 2, the issues of interest and the general framework to address them are defined through some template examples of the calibration literature. The statistical theory of PII is set up in Section 3. Section 4 is devoted to sequential extensions of PII and Section 5 concludes.

2. Calibration as econometrics of misspecified models

Our econometric formalization of calibration is introduced in this section through two leading examples. Firstly, the Mehra and Prescott (1985) approach to the equity premium puzzle provides a convenient example of the relevance of the PII framework for a statistical rationalization of calibration. Secondly, more involved issues like the role of exogenous variables and the usefulness of a sequential approach are described in a second subsection about more general DSGE empirical issues.

2.1. The equity premium puzzle

In their presentation of the calibration approach, Kydland and Prescott (1991) lays the emphasis on the crucial role of the research question which must be clearly defined.[1] Mehra and Prescott (1985) addresses the question whether the large differential between the average return on equity and average risk free interest rate can be accounted for by models neglecting any frictions in the Arrow and Debreu set up. The simple statement of this question defines on the one hand the structural parameters of interest and on the other hand the instrumental parameters through which the empirical evidence is summarized.

In order to statistically formalize the calibration concepts, we introduce in this section general notations that are consistently maintained herein.

First, the structural parameters of interest for Mehra and Prescott's question are two taste parameters of a representative agent: $\theta_1 = (\gamma, \alpha)'$ in a Lucas (1978) type consumption-based CAPM. The representative agent preferences over random consumption paths are described by a time-separable expected power utility function

$$E_0 \sum_{t=0}^{\infty} \gamma^t U(c_t),$$

where

$$U(c_t) = \frac{c_t^{1-\alpha} - 1}{1 - \alpha}$$

and c_t denotes the consumption at time t. Of course, this way of economically defining the structural parameters of interest is tightly linked to the economic setting the modeler has in mind and might be reducing since, while γ represents the subjective discount factor, α represents both relative risk aversion and inverse of the elasticity of intertemporal substitution. This implicitly assumes that this reduction has no incidence on the answer to the aforementioned question of interest. Anyway, we stress here that the structural parameters of interest θ_1 are intrinsically defined through economic paradigms rather than through falsifiable statistical relations.

Second, in this approach the structural model is empirically assessed through its ability to reproduce some stylized facts of interest like here the high value of the equity premium. In our statistical framework, these stylized facts are referred to as the set of instrumental

[1] Actually, the sole word question is used for a section title.

parameters denoted β. The empirical relevance of the structural model is assessed precisely through the matching between the observed instrumental characteristics and their theoretical counterparts consistent with the structural model.

Perhaps one of the most difficult issue for a close statement of the calibration methodology is that the reality check relies on additional assumptions which are not part of the economic theory of interest. These additional assumptions may require the specification of additional parameters θ_2 possibly of infinite dimension. In Mehra and Prescott (1985), these parameters θ_2 define the technology, that is the Markov chain assumed to govern the gross rate of dividend payments. More precisely, this gross rate x_t is described by a two states Markov chain:

$$Pr\{x_{t+1} = \lambda_j | x_t = \lambda_i\} = \phi_{ij}, \quad i,j = \{1,2\},$$

where

$$\lambda_1 = 1 + \mu + \delta, \quad \lambda_2 = 1 + \mu - \delta$$

and

$$\phi_{11} = \phi_{22} = \phi, \qquad \phi_{12} = \phi_{21} = 1 - \phi.$$

In other words $\theta_2 = (\mu, \delta, \phi)$. More generally, the vector θ of structural parameters is split into two parts θ_1 and θ_2 where θ_1 gathers the characteristics of interest while θ_2 corresponds to nuisance parameters which are needed for the statistical assessment. The most usual case is the one where θ_1 is related to preference specifications (taste parameters) and θ_2 describes environmental characteristics (technology parameters). However it may be the case that, as it is for the question above, one is not interested in a complete description of preferences. Then the specification of θ_1 focuses only on a subset of taste parameters (discount factor, risk aversion coefficient) while θ_2 may include other behavioral characteristics (e.g. elasticity of intertemporal substitution).

In any case, the main role of these nuisance parameters θ_2 consists in indexing a *binding function* between the structural parameters of interest θ_1 and the instrumental parameters β:

$$\beta = \widetilde{\beta}(\theta_1, \theta_2). \tag{1}$$

Of course, the value β of the instrumental parameters defined by (1) is the theoretical one and may not coincide with the (population) value of the observed one; this is precisely the question addressed by the calibration exercise. For sake of illustration, let us go into further details in the presentation of the Mehra and Prescott (1985) model. They show that the period return for the equity if the current state is i (with a level c_t of consumption) and the next period state is j is given by

$$r^{\mathrm{e}}_{ij} = \frac{\lambda_j (w_j + 1)}{w_i} - 1, \tag{2}$$

where w_1 and w_2 are computed from the Euler equation through the linear system of two equations:

$$w_i = \gamma \sum_{j=1}^{2} \phi_{ij} \lambda^{1-\alpha} (w_j + 1), \quad i = 1, 2.$$

In other words the expected return on the equity is

$$R^{\mathrm{e}} = \sum_{i,j=1}^{2} \pi_i \phi_{ij} r_{ij}^{\mathrm{e}}, \tag{3}$$

where $\pi = (\pi_1, \pi_2)'$ corresponds to the vector of stationary probabilities of the Markov chain. The same type of characterization is available for the risk free return R^{f} and omitted here.

Then formulas (2)–(3) define the binding function between $\theta_1 = (\gamma, \alpha)'$ and $\beta = g(R^{\mathrm{f}}, R^{\mathrm{e}}) = (R^{\mathrm{f}}, R^{\mathrm{e}} - R^{\mathrm{f}})'$ where the vector $g(\cdot)$ contains the moments of interest. Of course, this function is indexed by the additional parameters $\theta_2 = (\mu, \delta, \phi)'$ which characterize the Markov chain.

The specific feature of the calibration methodology with respect to more standard statistical inference appears precisely at this stage: since our goal is to ask whether, given the technology, there exist taste parameters capable of matching the returns data, this, according the Cechetti et al. (1993) "dictates that we proceed in two steps, first estimating the parameters of the endowment process, and then computing a confidence bound for the taste parameters γ and α".

With respect to more orthodox econometrics, this two steps procedure may arouse, at least, two types of criticism: First, even though the only parameters of interest are the taste parameters θ_1, one get in general more accurate estimators by a joint, possibly efficient, estimation of $\theta = (\theta_1', \theta_2')'$. Second, even when ignoring the efficiency issue, it is somewhat questionable with regard to consistent estimation to focus on taste parameters while the technology corresponds obviously to a caricature of the reality. Nobody may believe that the endowment process is conformable with a two states Markov chain and this misspecification presumably contaminates the estimation of the parameters of interest. In our opinion, a garbled answer to the above criticisms would consist in claiming that this procedure should not be regarded as an econometric one attempting to consistently estimate the parameters of interest. In this respect, we share Hansen and Heckman (1996) point of view that the distinction drawn between calibrating and estimating the parameters of a model is artificial at best.

Actually, the core principle of the calibration approach as illustrated in Mehra and Prescott paper's consists in concluding that the structural model is rejected on grounds of "computational experiments" leading to unlikely values of the parameters of interest. Namely, in Mehra and Prescott (1985) it is argued that computed values of the discount factor and the relative risk aversion parameter outside their commonly acknowledged range ($0 < \gamma < 1, 0 \leqslant \alpha \leqslant 10$) proves the misspecification of the structural model. How could they maintain such an argument if they did not think that these computed values are consistent estimators of something which makes sense?

Consequently, we think that calibration should also be interpreted in terms of consistent estimation of the parameters of interest, even though this issue is addressed in a nonstandard way in several respects:

- First, as explained above, it is often addressed in a negative way. The model is rejected because the estimators of its alleged parameters are obviously inconsistent.
- Second, consistency is the only focus of interest. Efficiency is irrelevant in this setting since the calibration exercises gather a huge amount of historical information such as

series of asset returns over the whole last century in such way that the efficient use of the information is not an issue at all.
- Third, calibrators are fully aware that consistency might fail, precisely due to the misspecification of the technology or more generally of the additional assumptions about the nuisance parameters θ_2. Indeed, fully cautious about that, they advocate calibration as a search for sensible values of θ_2.

The main goal of this paper is to statistically analyze into further details the latter point. To the extent that the aforementioned consistency requirement is maintained, the crucial concern is the following: When one uses the binding function $\widetilde{\beta}(\cdot, \bar{\theta}_2)$ indexed by a hypothetical value $\bar{\theta}_2$ of θ_2 to recover an estimate $\widehat{\theta}_1$ of the parameters of interest θ_1 from an empirical measurement $\widehat{\beta}$ of the instrumental parameters β by solving:[2]

$$\widehat{\beta} = \widetilde{\beta}(\widehat{\theta}_1, \bar{\theta}_2), \tag{4}$$

is there any hope that $\widehat{\theta}_1$ consistently estimates the true unknown value θ_1^0 of the structural parameters of interest? Before answering this question, three preliminary remarks are in order:

1. On the one hand, the sole idea of a true unknown value θ_1^0 of the structural parameters relies on the maintained hypothesis that the DGP is conformable to our structural ideas. This does not prevent from accounting for the calibrationnist approach which considers the estimation issue in a negative way as already explained.
2. On the other hand, we do not question here the consistency of the instrumental estimator $\widehat{\beta}$ since the instrumental parameters β^0 are essentially defined as the population value of $\widehat{\beta}$.
3. Finally, we consider for the moment that the binding function $\widetilde{\beta}(\cdot, \bar{\theta}_2)$, for any reasonable value $\bar{\theta}_2$, is well defined and known as is the case of the Mehra and Prescott (1985) framework. However, to capture complicated features of richer models, simulations at different levels of the forcing processes and parameters may be useful when analytical computation is intractable. This is perhaps the reason why calibrators have extensively used simulations.

The hope for getting a consistent estimator $\widehat{\theta}_1$ of θ_1^0 by solving (4) can then be supported by two alternative arguments according to our degree of optimism: Either, one adopts an optimistic approach wishing that history has provided sufficiently rich empirical evidence to determine without ambiguity a value $\bar{\theta}_2$ of the nuisance parameters. This is typically what is referred to as the calibration step. However, one should keep in mind that the technology is crudely misspecified (see the two states Markov chain above) in such a way that the estimator $\widehat{\theta}_1$ can be consistent only by chance whatever the choice of $\bar{\theta}_2$. Or, to be more cautious, one tries different values of $\bar{\theta}_2$ to check whether the outcome of the computational experiments is drastically changed. This is what is called the robustness of results in Mehra and Prescott (1985) and more generally the sensitivity analysis in the calibration literature.

[2] We do not mention the issue on overidentification which might prevent one from finding an exact solution to (4). See Section 3 for more details.

Of course, an ingenuous comment about this debate would be: one should jointly statistically estimate $(\beta, \theta_1, \theta_2)$ under the constraint (4). But this proposal is irrelevant in the calibration framework since the modeler knows a priori and before any statistical inference that the nuisance parameters θ_2 do not make sense on their own. Moreover, one of the main recommendations of this paper is to be suspicious in front of sophisticated strategies of model choice and fit about the technology characteristics. For instance, following Bonomo and Garcia (1994) it is true that by contrast with Cechetti et al. (1990) "a well-fitted equilibrium asset pricing model" may account for some stylized facts but one cannot be sure that the improvement in the technology specification is really relevant for the question of interest since misspecification is always guaranteed. For the same reason, one would not like neither to estimate the structural model through a large dimensional instrumental parameter β like a semi-nonparametric score (Bansal et al., 1995) nor assessing its goodness of fit with the associated large battery of diagnostic tests (Gallant et al. 1997; Tauchen et al., 1997).

Roughly speaking, all what really matters to validate the calibration exercise is that the structural model, when endowed with the pseudo-true value $(\theta_1^0, \bar{\theta}_2)$, encompasses the instrumental one in the sense that:

$$\beta^0 = \beta(\theta_1^0, \bar{\theta}_2).$$

We want to stress here that this encompassing condition is a sufficient condition for consistency of indirect estimators of the true unknown value $\theta_1^0$of the structural parameters of interest. This has almost nothing to do with the accuracy of the proxy of the technology provided by the nuisance parameters to the extent that the structural model is always "an abstraction of a complex reality" (Kydland and Prescott, 1991).

The calibration strategy adopted by Cechetti et al. (1993) reflects the concern for a parsimonious choice of the instrumental model given the technology process. These authors also investigate the equity premium through the first and second moments of the risk-free rate and the return to equity. As in Mehra and Prescott, the utility function is time-separable with a constant relative risk aversion. While Mehra and Prescott consider consumption and dividend as equal and then calibrate on an univariate Markov process, the model developed by Cechetti et al. (1993) explicitly disentangles consumption from dividends and the endowment process is defined by a bivariate consumption-dividends Markov-switching model.

Cechetti et al. (1993) are clearly aware of the problem of choosing a too large set of moments to estimate both structural parameters of interest and the endowment. They explicitly argue that it would not be well-suited to estimate the parameters of interest and the endowment process jointly by MLE procedure. Such an estimation strategy forces the model to match all the aspects of the data and it is unlikely that a simple model could reproduce adequately all those aspects. They formalize the calibration principle by the following two-step procedure. In a first step, the parameters of the endowment process are estimated through a subset of moments chosen to match the MLE estimates of a bivariate consumption-dividends Markov-switching model. In the second step, a confidence interval bound is computed for the taste parameters through first and second moments of returns data for a given endowment process. In our notation, this approach amounts to define two subvectors of instrumental parameters namely:

$$\beta_1 = \beta_1(\theta_1, \theta_2),$$
$$\beta_2 = \beta_2(\theta_2),$$

where $\theta_1 = (\gamma, \alpha)'$ and θ_2 gathers the parameters for the endowment process. The subvector $\beta_2(\cdot)$ corresponds to the subset of moments chosen to match the MLE estimates of the bivariate consumption-dividends Markov-switching model. The subvector $\beta_1(\cdot)$ contains the first and second moments of return data used to estimate the structural parameters θ_1 given the technology characterized by θ_2. As in Mehra and Prescott, the model evaluation relies on the plausibility of the confidence interval bound for the discount factor parameter (γ) and the relative risk aversion parameter (α). In other words, while the notation $\beta_2 = \beta_2(\theta_2)$ formalizes the fact that the calibrated values of θ_2 come themselves from some moment matching, and the notation $\beta_1 = \beta_1(\theta_1, \theta_2)$ stresses the fact that we are going to be parsimonious about the stylized facts the structural parameters θ_1 of interest are supposed to account for (for a given pseudo-true value of the nuisance structural parameters θ_2) it remains true that consistent confidence sets for θ_1 are going to be obtained from this two-step procedure only if one maintains the assumption of full encompassing of the instrumental model by the structural one.

2.2. General equilibrium approach to business cycles: an illustration

Kydland and Prescott (1982) introduced a neoclassical one-sector growth model driven by technology shocks to reproduce cyclical properties of US economy. The model includes a standard neoclassical production, standard preferences to describe agent's willingness to substitute intratemporally and intertemporally between consumption and leisure and an exogenously driven latent process characterizing technology shocks. The Kydland and Prescott's model and the subsequent macro dynamic equilibrium models based only on real shocks with no role for monetary shocks are called real business cycle (RBC) models.[3]

The clear-cut question addresses by Kydland and Prescott (1982) is the following: How much would the US economy have fluctuated if technology shocks had been the only source of fluctuations? Obviously the model is misspecified. In particular, it implies some unrealistic stochastic singularity for the vector of endogenous variables.[4]

This question addressed by Kydland and Prescott defines the moments (instrumental parameters) through which the empirical fit of the model has to be assessed. The instrumental parameters correspond to second moments describing the cyclical properties of US postwar economy. While these moments can be easily estimated from the data, simulations are often required to compute their theoretical counterpart. In the strategy advocated by Kydland and Prescott (1982) the answer to the question of interest is then given by an informal distance between empirical instrumental parameters and the instrumental parameters under the structural model. The values of the structural

[3]For extensions of this model see e.g. Hansen (1985), Beaudry and Guay (1996) and Burnside and Eichenbaum (1996).

[4]Some empirical applications bypass this misspecification problem by augmenting the theoretical solution of the model with a measurement error for each endogenous variables. The augmented model is then estimated by MLE (see Hansen and Sargent (1979), Christiano (1988)). See Watson (1993) and Ruge-Murcia (2003) for a discussion.

parameters are previously deduced from applied micro-studies or by matching long run properties of US economy.

For sake of illustration, we consider here a benchmark RBC model (King et al., 1988a, b). The social planner of this economy maximizes

$$E_0 \sum_{t=0}^{\infty} \gamma^t [\ln(C_t) + \phi \ln(L_t)],$$

where C_t is per capita consumption, L_t is leisure, γ is the discount factor and ϕ is the weight of leisure in the utility function. The intertemporal maximization problem is subject to the following budget constraint:

$$C_t + K_{t+1} - (1-\delta)K_t \leqslant K_t^{1-\alpha}(Z_t N_t)^{\alpha},$$

where K_t is the capital stock, N_t are the hours worked, Z_t is the labor augmenting technology process, α is the labor share in the Cobb–Douglas production function and δ the depreciation rate of the capital stock. As mentioned by Kydland and Prescott (1996), the law of motion of the latent process Z_t in the model is not provided by any economic theory. Additional assumptions which are neither given by economic theory nor by any statistical procedure are then required. Following King et al. (1988b), we consider here that the law of motion for Z_t is characterized by the following random walk with drift:

$$\ln Z_t = \mu + \ln Z_{t-1} + \varepsilon_t,$$

where μ is the growth rate of the economy and is ε_t i.i.d. Normal $(0, \sigma_\varepsilon)$. Obviously, this law of motion of the technology process is a caricature of the true unknown process. Consequently, this misspecification could presumably contaminate the estimation of the structural parameters of interest. However, with such a driven process, the log-linear solution of the model is compatible with a unit root process for output, consumption, investment and real wages (see King et al., 1988b, 1991) and cointegration relationships between these variables which are consistent with US data.

We consider here that there are four deep structural parameters in this model and three auxiliary parameters. In our notation, $\theta_1 = (\gamma, \delta, \alpha)'$ gathers the parameters of interest and $\theta_2 = (\phi, \mu, \sigma_\varepsilon)'$ the nuisance parameters needed for statistical implementation, that is to index the binding function. We will explain later why ϕ is considered as nuisance parameter.

While Mehra and Prescott ask the question of existence of reasonable values of parameters of interest able to reproduce the observed risk premium, the RBC modeler asks the question: "Given a set of parameters of interest calibrated from micro-evidence or long run averages, what is the ability of the model to reproduce some well documented "stylized facts"?"

As explained above, Mehra and Prescott (and Cechetti et al., 1993) considers estimation issue in a negative way: they search for values of structural parameters (θ_1 in our notation) reproducing as well as possible the observed instrumental parameters β. The goodness of fit of the model is assessed through the order of magnitude of these values. Kydland and Prescott (1982) evaluate the performance of the model by its ability to reproduce well defined "stylized facts" which are computed by simulations at given values of the structural parameters (θ_1). The assigned value of the parameter vector θ_1 comes from other applied studies or by matching long run average values for the economy. In contrast to Mehra and Prescott strategy, the instrumental parameters used to assess the model differs from the ones used to obtain an estimator of the structural parameters. More precisely, the strategy advanced by Kydland and Prescott (1982) consists in two steps:

- First, structural parameters are calibrated to values used in applied studies and to match long run average values.
- Second, the verification is implemented by judging the adequacy of the model to reproduce well chosen ‘‘stylized facts’’. When they could not find reliable estimations of a subset of parameters in economic literature or by matching long run properties, these parameters are treated as free parameters. Their values are then chosen to minimize the distance between the well chosen ‘‘stylized facts’’ of the US economy and the corresponding ones of the model.

The first step corresponding to calibration is the most controversial one. Indeed, several authors have shown that parameters obtained from micro-applied studies can be plugged to a representative agent model to produce empirically concordant aggregate model only under very special circumstances (see Hansen and Heckman (1996) for a discussion on this point). However, matching long run properties is more conformable to the estimation step in classical econometrics. In fact, this practice consists in matching a just-identified set of moments where the corresponding instrumental parameters are the long-run averages. For instance, Kydland and Prescott (1982) calibrate the deterministic version of their model so that consumption/investment shares, factor/income shares, capital/output ratio, leisure/market-time shares and depreciation shares match the average values of US economy. Since this matching is not done through a formal GMM-type estimation procedure,[5] uncertainty inherent to computed values is not taken into account in the results.

The matching of long run properties of the economy corresponds in our setting to obtaining an estimator of $\theta = (\theta_1', \theta_2')'$ by matching

$$\beta = \beta(\theta_1, \theta_2),$$

where β captures these long run average properties. It is important to note that these long run properties correspond to stationary transformations of variables. Generally speaking, to avoid an obvious violation of common theoretical assumptions as listed in Section 3 below, moments conditions β must be defined on stationary transformations of variables.[6]

Moreover, the notion of ‘‘free parameters’’ in Kydland and Prescott (1982) corresponds to the idea that only a subset θ_{21} of nuisance structural parameters θ_2 actually shows up in the binding function $\beta(\theta_1, \theta_2)$. More precisely there is what we are going to define as *partial encompassing* of the instrumental model by the structural one because θ_2 is partitioned in $\theta_2 = (\theta_{21}, \theta_{22})$ and $\beta^0 = \beta(\theta_1^0, \bar{\theta}_{21})$ for some pseudo-true value $\bar{\theta}_{21}$ of θ_{21}. Of course finite sample and possibly simulation-based counterparts of $\beta(\theta_1^0, \bar{\theta}_{21})$ may depend on some given value $\bar{\theta}_{22}$ of θ_{22} needed to characterize the DGP but, as far as asymptotic estimation is concerned, the pseudo-true value $(\theta_1^0, \bar{\theta}_{21})$ obtained by matching with the instrumental parameters do not depend on θ_{22}. This invariance comes at the price of not identifying the free parameters θ_{22} from the instrumental parameters β.

[5]See Christiano and Eichenbaum (1992) and Burnside and Eichenbaum (1996) for the estimation of structural parameters by a just-identified GMM.

[6]See Christiano and Eichenbaum (1992), Burnside and Eichenbaum (1996), Beaudry and Guay (1996), among others, for the estimation of structural models with technology process characterized by a random walk with drift.

The verification step (second step) performed by the calibrator precisely involves the identification of these free parameters. While for calibrators it is based on a quite informal distance criterion for selected "stylized facts", it can be formalized in our setting by a choice of additional instrumental parameters corresponding to the "stylized facts" to reproduce. In fact, we try to judge whether we can reject with a certain metric the following null hypothesis:

$$\psi = \psi(\theta_1^0, \bar{\theta}_{21}, \bar{\theta}_{22})$$

evaluated at the pseudo-true value obtained $(\theta_1^0, \bar{\theta}_2)$ with the instrumental parameters $\mathcal{N}_\psi$ for some convenient choice of the pseudo-true value $\bar{\theta}_{22}$. In other words, the estimator of the pseudo-true value $\bar{\theta}_{22}$ of the nuisance parameters is obtained from the sample (and possibly simulation based) counterpart of some minimum distance program:

$$\bar{\theta}_{22} = \arg \min_{\theta_{22} \in \Theta_{22}} (\psi^0 - \widetilde{\psi}(\theta_1^0, \bar{\theta}_{21}, \theta_{22}))' \Omega^\psi (\psi^0 - \widetilde{\psi}(\theta_1^0, \bar{\theta}_{21}, \theta_{22})), \tag{5}$$

where Ω^ψ is a positive matrix on R^{q_ψ} and $q_\psi = \dim \psi$. For the benchmark RBC model, the free parameter ϕ corresponding to the weight of leisure in the utility function may be difficult to estimate at the first step. In such a situation, an estimator can then be obtained by (5). In a more complicated model, Kydland and Prescott fix seven parameters by minimizing the distance between the model and data for 23 moments describing US business cycle. Those parameters are the substitutability of inventories and capital, two parameters determining intertemporal substitutability of leisure, the risk aversion parameter and three parameters for the technology process.

It is worth noticing that, even when more recent econometric studies of DSGE models claim to use more formal statistical techniques because "an important advantage of our GMM procedures, however, is that they let us quantify the degree of uncertainty in our estimates of the model's parameters" (Christiano and Eichenbaum, 1992), they often realize that efficient GMM a la Hansen (1982) is not well-suited in such a misspecified setting and what they actually do resembles much more the two steps of calibration and verification described above. For sake of illustration, we focus here on the example of Christiano and Eichenbaum (1992) but the general features of this study as described below are largely shared with most empirical studies of DSGE models. Roughly speaking, their application of GMM is far to be orthodox for the following reasons.

First, as above, there is an implicit partition of the vector θ of structural parameters between three subsets. The vector θ_1 of structural parameters of interest includes α, δ, ϕ in order to characterize both the central planner utility function (up to the subjective discount factor γ) and the production function. Actually, at it is written, the utility function includes another parameter N which is the time endowment of the representative consumer seen as an upper bound of the leisure time L_t defined above. It is rather clear that the two nuisance structural parameters $\theta_{21} = (\gamma, N)$ are highly difficult to estimate but however are "key parameters" that may contaminate the estimation of the structural parameters θ_1 of interest. Therefore, in spite of the claim of taking parameters uncertainty into account, the value of these two parameters is fixed only from the calibrator knowledge, without any attempt to identify them statistically, while strictly speaking, they are identified by the structural model. In other words, a very parsimonious binding finction $\beta^0 = \beta(\theta_1, \bar{\theta}_{21})$, actually just identified with respect to θ_1 is written to get partial encompassing for the fixed value $\bar{\theta}_{21}$ of the nuisance parameters. To do so, the authors neglect on purpose the identifying content (about both θ_1 and θ_{21}) of Euler intertemporal optimality conditions and only use unconditional moment restrictions to just identify θ_1.

Second, while the Christiano and Eichenbaum (1992) model includes five other structural parameters characterizing the dynamics of the aggregate shock to technology and of an additional process of public consumption as well, these parameters define a vector θ_{22} which may not be simultaneously estimated with θ_1. Actually, Christiano and Eichenbaum (1992) do estimate them simultaneously, but only within a just identified GMM framework where the estimation of θ_{22} does not contaminate the estimation of the structural parameters of interest θ_1. It is important to realize that in more involved DSGE models with latent processes where structural moments need to be simulated because there are no close form formulas, this kind of cut between the two sets of parameters could be obtained only by giving up a number of possible binding functions between structural parameters and observable moments. As a matter of fact, such a problematic interaction between parameters of interest and nuisance parameters is only met by Christiano and Eichenbaum (1992) when they come to the second step of model verification about "labor-market moments". Then, as explained above, is only when it comes to the assessment of the ability of the model to reproduce some stylized facts that the interaction between nuisance parameters θ_{22} and parameters of interest θ_1 is explicitly acknowledged. The latter stylized facts define an additional set $\psi = \psi(\theta_1^0, \bar{\theta}_{21}, \bar{\theta}_{22})$ of moments conditions which are not used for a joint efficient estimation of all the structural parameters within the first step of calibration.

Finally, it is worth reminding that, while it has been possible in this simple model to perform GMM inference from moment conditions that are available in closed form, it is often the case that preliminary simulations in the structural model are needed to get a simulation-based counterpart of the binding function. In such a case, it would be a pity to contaminate the estimation of the structural parameters of interest, simply because a simulator of exogenous variables, like public consumption in Christiano and Eichenbaum (1992) model, has been wrongly specified. In other words, common formulas of simulated method of moments kind of inference have to be corrected for the effect of not simulating the paths of observed exogenous processes.

3. A partial indirect inference approach to calibration

We present in this section the principles of PII. While this methodology may be relevant for other applications (see e.g. Dridi and Renault (1998) for an application to stochastic volatility models), we give here the general theory, as an extension of Gouriéroux et al. (1993) II, but we set the focus on the calibration framework as formalized in Section 2.

The main goal of this section is to give a precise content to the calibration kind of interpretation of II, as put forward in Section 2, that is "given that the model is false", some elements of truth involved in the model (for instance some taste parameters) should be caught by matching some well-chosen moments. The rigorous meaning of "elements of truth" lies in the semi-parametric modeling widely adopted in modern econometrics as an alternative to the hopeless search for a well-specified parametric model that is more often than not impossible to deduce from economic theory. On the opposite, the partial approach to II specifies only some parameters of interest raised out by the underlying economic theory. We first present the theoretical results (consistency, asymptotic probability distribution) available for PII.

3.1. The general framework

The data consist in the observation of a stochastic process $\{(y_t, x_t), t \in Z\}$ for $t = 1, \ldots, T$. We denote by P_0 the true unknown p.d.f. of $\{(y_t, x_t), t \in Z\}$.

Assumption (A1). (i) P_0 belongs to a family $\mathscr{P}$ of p.d.f. on $(\mathscr{X} \times \mathscr{Y})^Z$.

(ii) $\widetilde{\theta}_1$ is an application from $\mathscr{P}$ onto a part $\Theta_1 = \widetilde{\theta}_1(\mathscr{P})$ of R^{p_1}.

(iii) $\widetilde{\theta}_1(\mathscr{P}_0) = \theta_1^0$, the true unknown value of the parameters of interest, belongs to the interior $\overset{\circ}{\Theta}_1$ of Θ_1.

$\widetilde{\theta}_1(\mathscr{P}) = \theta_1$ is the vector unknown parameters of interest. Typically, in the case of a stationary process $\{(y_t, x_t), t \in Z\}$, it may be defined through a set h of identifying moment restrictions:

$$\underset{P}{E}\, h(y_t, x_t, y_{t-1}, x_{t-1}, \ldots, y_{t-K}, x_{t-K}, \theta_1) = 0 \quad \Longrightarrow \quad \theta_1 = \widetilde{\theta}_1(P).$$

In such a partially parametric model, not only the MLE estimator is no longer available, but even more robust M-estimators or minimum distance estimators may be intractable due to a complicated dynamic structure of $\mathscr{P}$. This is the reason why we refer to II associated with a given pair of "structural" model (used as simulator) and "auxiliary" (or "instrumental") criterion.

In order to get a simulator useful for PII on θ_1, we plug the partially parametric model defined by (A1) into a structural model that is fully parametric and misspecified in general since it introduces additional assumptions on the law of motion of (y, x) which are not suggested by any economic theory. These additional assumptions require a vector θ_2 of additional parameters. The vector θ of "structural parameters" is thus given by $\theta = (\theta_1', \theta_2')'$. We then formulate a nominal assumption (B1) to specify a structural model conformable to the previous section, even though we know that (B1) is likely to be inconsistent with the true DGP.[7]

Nominal assumptions (B1). $\{(y_t, x_t), t \in Z\}$ is a stationary process conformable to the following nonlinear simultaneous model:

- $$r(y_t, y_{t-1}, x_t, u_t, \theta) = 0,$$ $$\varphi(u_t, u_{t-1}, \varepsilon_t, \theta) = 0$$
  $\theta = (\theta_1', \theta_2')' \in (\Theta_1 \times \Theta_2) = \Theta$ a compact subset of $R^{p_1+p_2}$,
- the exogenous process $\{x_t, t \in Z\}$ is independent of $\{\varepsilon_t, t \in Z\}$,
- $\{\varepsilon_t, t \in Z\}$ is a white noise with a known distribution G_*.

Then, for each given value of the parameters θ, it is possible to simulate values $\{\widetilde{y}_1(\theta, z_0), \ldots, \widetilde{y}_T(\theta, z_0)\}$ conditionally on the observed path of exogenous variables $\{x_1, \ldots, x_T\}$ and for given initial conditions $z_0 = (y_0, u_0)$. This is done by simulating values $\{\widetilde{\varepsilon}_1, \ldots, \widetilde{\varepsilon}_T\}$ from G_*. We denote by P_* the probability distribution of the process $\{x_t, \varepsilon_t, t \in Z\}$.

[7] We denote by B the nominal assumptions, that is assumptions that are used for a quasi-II (by extension of the QMLE).

We focus here on II about the true value θ_1^0 of the parameters of interest θ_1. The II principle is still defined from the two basic components: a "structural" model (B1) and an instrumental model $\mathcal{N}_\beta$ which defines pseudo-true values of instrumental parameters as limits in probability of some extremum estimator associated to a criterion function $Q_T(\underline{y}_T, \underline{x}_T, \beta)$ to minimize. Here we assume for simplicity that $\beta \in B$ a compact subset of R^q and $\underline{y}_T$, $\underline{x}_T$ denote the lagged values of y_t and x_t for a fixed number K of lags.

For example, in the case of moment conditions, the corresponding criterion is defined as

$$Q_T(\underline{y}_T, \underline{x}_T, \beta) = \frac{1}{2}\left(\frac{1}{T}\sum_{t=1}^{T} g(\underline{y}_t, \underline{x}_t) - \beta\right)'\left(\frac{1}{T}\sum_{t=1}^{T} g(\underline{y}_t, \underline{x}_t) - \beta\right).$$

We introduce the estimators $\widehat{\beta}_T$ and $\widetilde{\beta}_{TS}(\theta_1, \theta_2)$ associated with the general criterion:

$$\widehat{\beta}_T = \arg\min_{\beta\in B} Q_T(\underline{y}_T, \underline{x}_T, \beta),$$

$$\widetilde{\beta}^s_T(\theta_1, \theta_2) = \arg\min_{\beta\in B} Q_T(\underline{\widetilde{y}}^s_T(\theta, z_0^s), \underline{x}_T, \beta),$$

$$\widetilde{\beta}_{TS}(\theta_1, \theta_2) = \frac{1}{S}\sum_{s=1}^{S} \widetilde{\beta}^s_T(\theta_1, \theta_2),$$

where $\underline{\widetilde{y}}^s_t(\theta, z_0^s) = \{\widetilde{y}^s_t(\theta, z_0^s), \widetilde{y}^s_{t-1}(\theta, z_0^s), \ldots, \widetilde{y}^s_{t-K}(\theta, z_0^s)\}$ for S simulated paths $s = 1, 2, \ldots, S$ associated to a given value $\theta = (\theta_1', \theta_2')'$ of the structural parameters. Note that by contrast with a current practice in calibration studies, the simulation noise can be reduced by simulating S paths of length T for the endogenous variables (while repeating S times the observed bias of exogenous variables) with T equal to the length of observed paths and S larger than one. Roughly speaking, when the simulator is well-specified, this will lead to multiply the standard asymptotic variance matrices of GMM by a factor of (1 + 1/S) to take into account the simulation noise for simulated moments. With a misspecified structural model used as a simulator, the factor (1 + 1/S) does not show up anymore since observed moments and simulated moments are no longer produced from the same DGP. However, as shown in asymptotic variance formulas below, increasing S still provides variance reduction like (1/S). By contrast, it is important to simulate paths of length T to get the same order of magnitude of finite sample bias in estimators of instrumental parameters computed on both observed and simulated paths. The key point is that, such a finite sample path, while especially significant for estimating persistence of dynamic processes, will be erased by matching observed and simulated moments which involve the same kind of finite sample bias (see Gouriéroux et al., 2000).

Under usual regularity conditions, estimators computed on observed paths and simulated paths as well converge, when T is going to infinity, uniformly in (θ_1, θ_2) to

$$P_0 \lim_{T\to\infty} \widehat{\beta}_T = \beta^0 = \beta(P_0),$$

$$P_* \lim_{T\to\infty} \widetilde{\beta}_{TS} = \widetilde{\beta}(\theta_1, \theta_2).$$

We refer to $P_0 \lim_{T\to+\infty}$ and $P_* \lim_{T\to+\infty}$ as the limit with respect to the P_0 and the P_* probabilities when T goes to infinity.

Assumption (A2). $\widetilde{\beta}(\cdot,\cdot)$ is one-to-one.

According to Gouriéroux and Monfort (1995) terminology, β^0 is the true value of instrumental parameters and $\widetilde{\beta}(\cdot,\cdot)$ is the binding function from the structural model to the instrumental one.

A PII estimators $\widehat{\theta}_{1,TS}$ is then defined as follows:

$$\widehat{\theta}_{TS} = (\widehat{\theta}'_{1,TS}, \widehat{\theta}'_{2,TS})' = \arg \min_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2} [\widehat{\beta}_T - \widetilde{\beta}_{TS}(\theta_1,\theta_2)]'\widehat{\Omega}_T[\widehat{\beta}_T - \widetilde{\beta}_{TS}(\theta_1,\theta_2)],$$

where $P_* \lim_{T\to+\infty} \widehat{\Omega}_T = \Omega$ is positive definite matrix on R^q.

In order to derive a necessary and sufficient condition for the consistency of the PII estimator $\widehat{\theta}_{1,TS}$ to θ_1^0, we define the so-called "generalized inverse" $\widetilde{\beta}^-$ of $\widetilde{\beta}$ by

$$\widetilde{\beta}^-(\beta) = \arg \min_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2} \|\beta - \widetilde{\beta}(\theta_1,\theta_2)\|_\Omega.$$

In our partially parametric setting, we are only interested in the projection of $\widetilde{\beta}^-[\beta(P)]$ on the set Θ_1 of the parameters of interest. Let us denote by Q_1 the projection operator:

$$\begin{aligned} Q_1 : R^{p_1} \times R^{p_2} &\to R^{p_1}, \\ (\theta'_1, \theta'_2)' &\to \theta_1. \end{aligned}$$

We are then led to the following consistency criterion:

Proposition 3.1. *Under assumptions* (A1)–(A2), $\widehat{\theta}_{1,TS}$ *is a consistent estimator of the parameters of interest* θ_1^0 *if and only if, for any* P *in the family* $\mathscr{P}$ *of p.d.f. delineated by the model* (A1)

$$Q_1[\widetilde{\beta}^-(\beta(P))] = \widetilde{\theta}_1(P).$$

In order to test the consistency property, we focus on a sufficient encompassing condition. We say that (B1) endowed with the pseudo-true value $(\theta_1^{0\prime}, \bar{\theta}'_2)'$ fully encompasses $(\mathscr{N}_\beta)$ if:

$$\beta^0 = \widetilde{\beta}(\theta_1^0, \bar{\theta}_2).$$

In this framework, we get easily under standard regularity conditions a sufficient condition for the consistency of the PII estimator $\widehat{\theta}_{1,TS}$:

Proposition 3.2. *Under assumptions* (A1)–(A2), *if there exists* $\bar{\theta}_2 \in \Theta_2$ *such that* (B1) *endowed with the pseudo-true value* $(\theta_1^{0\prime}, \bar{\theta}'_2)'$ *fully encompasses* $(\mathscr{N}_\beta)$, *then* $\widehat{\theta}_{1,TS}$ *is a consistent estimator of the parameters of interest* θ_1^0.

Proposition 3.2 is a direct corrollary of Proposition 3.1.

When the structural misspecified model (B1) endowed with the pseudo-true value $(\theta_1^{0\prime}, \bar{\theta}'_2)'$ for $\bar{\theta}_2 \in \Theta_2$ does not fully encompasses the instrumental model $\mathscr{N}_\beta$, we are led to extend the encompassing concept to a property of partial encompassing defined through a subvector β_1^0 of q_1 instrumental parameters ($q_1 \leqslant q$). The corresponding subvector function $\widetilde{\beta}_1(\cdot,\cdot)$ of the binding function is defined from $\Theta_1 \times \Theta_{21}$ onto R^{q_1}:

$$\widetilde{\beta}_1 : \Theta_1 \times \Theta_{21} \to R^{q_1}, \tag{6}$$

$$(\theta_1, \theta_{21}) \to \tilde{\beta}_1(\theta_1, \theta_{21}), \tag{7}$$

where θ_{21} corresponds to the subvector of the nuisance parameters $\theta_2 = (\theta'_{21}, \theta'_{22})'$ which does play a role in the first q_1 components of the binding function β. θ_{21} belongs to Θ_{21}, subset of $R^{p_{21}}$ with the assumed factorization of the nuisance parameters set of $\Theta_2 = \Theta_{21} \times \Theta_{22}$. We say that (B1) endowed with the pseudo-true value $(\theta_1^{0\prime}, \bar{\theta}'_2)'$ partially encompasses $\mathcal{N}_\beta$ if the following conditions are fulfilled:

(i) $\tilde{\beta}_1(\cdot, \cdot)$ is one-to-one,
(ii) $\beta_1^0 = \tilde{\beta}_1(\theta_1^0, \bar{\theta}_{21})$.

We introduce the following estimators $\hat{\beta}_{1,T}$, and $\tilde{\beta}^s_{1,T}(\theta_1, \theta_2)$, respectively, defined as the subvectors of size q_1 of the estimators $\hat{\beta}_T$ and $\tilde{\beta}_{TS}(\theta_1, \theta_2)$. These estimators converge uniformly in θ_1, θ_2 to:

$$P_0 \lim_{T\to\infty} \hat{\beta}_{1,T} = \beta_1^0,$$
$$P_* \lim_{T\to\infty} \tilde{\beta}_{1,TS}(\theta_1, \theta_2) = \tilde{\beta}_1(\theta_1, \theta_{21}).$$

In this context, since the PII estimator $\hat{\theta}_{1,TS}$ is possibly not consistent for θ_1^0, we propose to focus on another class of partial indirect estimator $\hat{\theta}_{1,TS}(\bar{\theta}_{22})$ based on a subvector β_1 of the instrumental parameters and defined by:

$$\begin{aligned}\hat{\theta}^1_{TS}(\bar{\theta}_{22}) &= (\hat{\theta}^{1\prime}_{1,TS}(\bar{\theta}_{22}), \hat{\theta}^{1\prime}_{21,TS}(\bar{\theta}_{22}))' \\ &= \arg \min_{(\theta_1,\theta_{21})\in\Theta_1\times\Theta_{21}} [\hat{\beta}_{1,T} - \tilde{\beta}_{1,TS}(\theta_1, \theta_{21}, \bar{\theta}_{22})]'\hat{\Omega}_{1,T}[\hat{\beta}_{1,T} - \tilde{\beta}_{1,TS}(\theta_1, \theta_{21}, \bar{\theta}_{22})],\end{aligned}$$

where $P_* \lim_{T\leftarrow+\infty} \hat{\Omega}_{1,T} = \Omega_1$ is a positive definite matrix. We denote by $\bar{\theta}_{22}$ the value assigned to the nuisance parameters θ_{22} in order to perform the simulations. In this framework, we are able to prove the following sufficient condition for the consistency of the PII estimator $\hat{\theta}^1_{1,TS}(\bar{\theta}_{22})$:

Proposition 3.3. *Under assumptions* (A1)–(A2), *and if there exists* $\bar{\theta}_2 \in \Theta_2$ *such that* (B1) *endowed with the pseudo-true value* $(\theta_1^{0\prime}, \bar{\theta}'_2)'$ *partially encompasses* $\mathcal{N}_\beta$, *then* $\hat{\theta}^1_{1,TS}(\bar{\theta}_{22})$ *is a consistent estimator of the parameters of interest* θ_1^0.

3.2. Asymptotic probability distribution of partial indirect inference estimators

In this section, we derive the asymptotic probability distribution of PII estimators. While the general statement of regularity conditions and the methodology of proofs remain conformable to standard asymptotic theory of extremum estimation, two specific features have to be emphasized. First the possible discrepancies between the DGP and the simulator lead to define two sets of information kind of matrices, whatever their expression as gradient of the score or Hessian of the objective function. Second, the fact same some exogenous variables are not simulated but only duplicated may introduce some perverse correlations between observed and simulated paths. We distinguish two sets of asymptotic results, depending upon the fact they are obtained under the maintained assumption of either full or only partial encompassing.

3.2.1. Full-encompassing partial indirect inference estimator

We focus here on the asymptotic properties of the II estimator $\widehat{\theta}_{TS}$ under the full-encompassing hypothesis: there exists $\bar{\theta}_2 \in \Theta_2$ such that (B1) endowed with the pseudo-true value $(\theta_1^{0\prime}, \bar{\theta}_2')'$ fully encompasses $\mathcal{N}_\beta$. Let us then assume as usual for extremum estimation that:

First,

$$\text{(A3)} \quad \frac{1}{\sqrt{T}} \frac{\partial Q_T}{\partial \beta} (\underline{y}_T, \underline{x}_T, \beta^0),$$

is asymptotically normally distributed with mean zero and with an asymptotic covariance matrix I_0 and second:

$$\text{(A4)} \quad J_0 = P_0 \lim_{T\to\infty} \frac{\partial^2 Q_T}{\partial\beta\partial\beta'} (\underline{y}_T, \underline{x}_T, \beta^0).$$

Of course, since Q_T may not be the log-likelihood, the information matrix equality is not guaranteed and I_0 and J_0 may differ. Moreover, since DGP and simulator may differ, we must also consider a second set of similar matrices associated to the simulator when the pseudo-true value of the parameters is used for simulation. More precisely, we assume that:

$$\text{(A5)} \quad \frac{1}{\sqrt{T}} \frac{\partial Q_T}{\partial \beta} (\underline{\tilde{y}}_T^s(\theta_1^0, \bar{\theta}_2', z_0^s), \beta^0),$$

is asymptotically normally distributed with mean zero and with an asymptotic covariance matrix I_0^* and independent of the initial values $z_0^s, s = 1, \ldots, S$, and

$$\text{(A6)} \quad J_0^* = P_* \lim_{T\to\infty} \frac{\partial^2 Q_T}{\partial\beta\partial\beta'} (\underline{\tilde{y}}_T^s(\theta_1^0, \bar{\theta}_2', z_0^s), \beta^0).$$

Note that this setting is also relevant for performing simulated MLE, as it is popular nowadays for estimating DSGE models. While we have known since White (1982) that, as soon as one think about possible misspecifications of the likelihood functions, like omitted variables, omitted heteroskedasticity, and so on, one should use a robustified asymptotic covariance matrix taking into account that I_0 and J_0 may differ, we stress here that it is as important to realize that I_0, J_0, I_0^* and J_0^* may be four different matrices. The four ones are going to be at stake in asymptotic variances of simulation-based estimators.

Finally, one must also take into account the perverse correlation between observed and simulated paths (and associated score functions) due to the fact that observed exogenous variables paths are duplicated within simulated paths, in order to avoid a likely wrong simulation of exogenous variables. Once more, this is even more relevant in the case of simulated likelihood approaches where no such thing that a probability distribution for exogenous processes should be specified. In this case too, one must distinguish score functions as produced by the true DGP and score functions produced by the simulator, so that two matrices of asymptotic covariance have to be defined:

$$\text{(A7)}\quad \lim_{T\to+\infty} \operatorname{Cov}_*\left\{\frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{y}_T, \underline{x}_T, \beta^0), \frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\widetilde{y}}^s_T(\theta^0_1, \bar{\theta}'_2, z^s_0), \beta^0)\right\} = K_0,$$

independent of the initial values $z^s_0, s = 1, \ldots, S$, and

$$\text{(A8)}\quad \lim_{T\to+\infty} \operatorname{Cov}_*\left\{\frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\widetilde{y}}^s_T(\theta^0_1, \bar{\theta}'_2, z^s_0), \beta^0), \frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\widetilde{y}}^l_T(\theta^0_1, \bar{\theta}'_2, z^l_0), \beta^0)\right\} = K^*_0,$$

independent of the initial values z^s_0 and z^l_0, for $s \neq \ell$.

Finally, as for standard minimum distance estimation asymptotic theory, we assume that:

$$\text{(A9)}\quad P_* \lim_{T\to+\infty} \frac{\partial\widetilde{\beta}^s_T}{\partial\theta'}(\theta^0_1, \bar{\theta}_2) = \frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta^0_1, \bar{\theta}_2),$$

is full-column rank (p).

We are now able to state the following result:

Proposition 3.4. *Under the null hypothesis of full encompassing and assumptions* (A1)–(A9), *the optimal II estimator $\widehat{\theta}^*_{TS}$ is obtained with the weighting matrix Ω^* defined below. It is asymptotically normal, when S is fixed and T goes at infinity*:

$$\sqrt{T}\begin{pmatrix}\widehat{\theta}_{1,TS} - \theta^0_1 \\ \widehat{\theta}_{2,TS} - \bar{\theta}_2\end{pmatrix} \xrightarrow{\text{D}} \mathcal{N}(0, W(S, \Omega))$$

with

$$W(S, \Omega) = \left\{\frac{\partial(\widetilde{\beta})'}{\partial\theta}(\theta^0_1, \bar{\theta}_2)(\Phi^*_0(S))^{-1}\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta^0_1, \bar{\theta}_2)\right\}^{-1}, \tag{8}$$

$$\begin{aligned}\Omega^* &= \Phi^*_0(S)^{-1},\\ \Phi^*_0(S) &= J_0^{-1}I_0J_0^{-1} + \frac{1}{S}J_0^{*-1}I^*_0J_0^{*-1} + \left(1 - \frac{1}{S}\right)J_0^{*-1}K^*_0J_0^{*-1}\\ &\quad - J_0^{-1}K_0J_0^{*-1} - J_0^{*-1}K'_0J_0^{-1}.\end{aligned} \tag{9}$$

Proof. see Appendix A. □

Note that in the case where the structural model (B1) is well specified, $\Phi^*_0(S)$ reduces to $(1 + \frac{1}{S})J_0^{-1}(I_0 - K_0)J_0^{-1}$ since then $K_0 = K'_0$.

3.2.2. Partial-encompassing partial indirect inference estimator

We now focus on the asymptotic properties of the II estimator $\widehat{\theta}^1_{TS}(\bar{\theta}_{22})$ under the partial encompassing hypothesis $H^1_0(\bar{\theta}_{22})$. The main difference with the full encompassing case is that now all the asymptotic covariances depend on some $\bar{\theta}_{22}$ which has been calibrated before any formal estimation procedure. We first maintain assumptions (A3) and (A4) and we denote $\widetilde{\beta}^0(\bar{\theta}_{22}) = \widetilde{\beta}(\theta^0_1, \bar{\theta}_2)$ for the given value $\bar{\theta}_{22}$ of the nuisance parameters. We made the following assumptions for a given value $\bar{\theta}_{22}$:

$$\text{(A10)}\quad \lim_{T\to+\infty} \mathrm{Cov}_*\left\{\frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{y}_T, \underline{x}_T, \beta^0), \frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\tilde{y}}^s_T(\theta^0_1, \bar{\theta}'_2, z^s_0), \tilde{\beta}^0(\bar{\theta}_{22}))\right\} = K_0(\bar{\theta}_{22}),$$

independent of the initial values $z^s_0, s = 1, \ldots, S$.

$$\text{(A11)}\quad \frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\tilde{y}}^s_T(\theta^0_1, \bar{\theta}'_2, z^s_0), \tilde{\beta}^0(\bar{\theta}_{22})),$$

is asymptotically normally distributed with mean zero and with an asymptotic covariance matrix $I^*_0(\bar{\theta}_{22})$ and independent of the initial values $z^s_0, s = 1, \ldots, S$.

$$\text{(A12)}\quad J^*_0(\bar{\theta}_{22}) = P_* \lim_{T\to\infty} \frac{\partial^2 Q_T}{\partial\beta\partial\beta'}(\underline{\tilde{y}}^s_T(\theta^0_1, \bar{\theta}'_2, z^s_0), \tilde{\beta}^0(\bar{\theta}_{22})),$$

$$\text{(A13)}\quad \lim_{T\to+\infty} \mathrm{Cov}_*\left\{\frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\tilde{y}}^s_T(\theta^0_1, \bar{\theta}'_2, z^s_0), \tilde{\beta}^0(\bar{\theta}_{22})), \frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\tilde{y}}^l_T(\theta^0_1, \bar{\theta}'_2, z^l_0), \tilde{\beta}^0(\bar{\theta}_{22}))\right\} = K^*_0(\bar{\theta}_{22}),$$

independent of the initial values z^s_0 and z^l_0, for $s \neq \ell$.

$$\text{(A14)}\quad P_* \lim_{T\to+\infty} \frac{\partial\tilde{\beta}^s_{1,T}}{\partial\begin{pmatrix}\theta_1\\ \theta_{21}\end{pmatrix}'}(\theta^0_1, \bar{\theta}_2) = \frac{\partial\tilde{\beta}_1}{\partial\begin{pmatrix}\theta_1\\ \theta_{21}\end{pmatrix}'}(\theta^0_1, \bar{\theta}_{21}),$$

is full-column rank $(p_1 + p_{21})$. We are then able to prove the following result:

Proposition 3.5. *Under the null hypothesis* $H^1_0(\bar{\theta}_{22})$, *assumptions* (A1)–(A4), (A10)–(A14), *the optimal II estimator* $\hat{\theta}^{1*}_{TS}(\bar{\theta}_{22})$ *is obtained with the weighting matrix* $\Omega^*_1(\bar{\theta}_{22})$ *defined below. It is asymptotically normal, when S is fixed and T goes to infinity*:

$$\sqrt{T}\begin{pmatrix}\hat{\theta}^1_{1,TS}(\bar{\theta}_{22}) - \theta^0_1\\ \hat{\theta}^1_{21,TS}(\bar{\theta}_{22}) - \bar{\theta}_{21}\end{pmatrix} \xrightarrow{D} \mathcal{N}\left(0, W_1(S, \Omega^*_1(\bar{\theta}_{22}))\right)$$

with

$$W_1(S, \Omega^*_1(\bar{\theta}_{22})) = \left[\frac{\partial\tilde{\beta}'_1}{\partial\begin{pmatrix}\theta_1\\ \theta_{21}\end{pmatrix}}(\theta_1, \bar{\theta}_{21})(\Phi^*_{0,1}(S, \bar{\theta}_{22}))^{-1}\frac{\partial\tilde{\beta}_1}{\partial\begin{pmatrix}\theta_1\\ \theta_{21}\end{pmatrix}'}(\theta^0_1, \bar{\theta}_{21})\right]^{-1},$$

$$\Omega^*_1(\bar{\theta}_{22}) = \Phi^*_{0,1}(S, \bar{\theta}_{22})^{-1},$$

$$\begin{aligned}\Phi^*_0(S) = {} & J^{-1}_0 I_0 J^{-1}_0 + \frac{1}{S}J^{*-1}_0(\bar{\theta}_{22})I^*_0(\bar{\theta}_{22})J^{*-1}_0(\bar{\theta}_{22})\\ & + \left(1 - \frac{1}{S}\right)J^{*-1}_0(\bar{\theta}_{22})K^*_0(\bar{\theta}_{22})J^{*-1}_0(\bar{\theta}_{22})\\ & - J^{-1}_0 K_0(\bar{\theta}_{22})J^{*-1}_0(\bar{\theta}_{22}) - J^{*-1}_0(\bar{\theta}_{22})K'_0(\bar{\theta}_{22})J^{-1}_0.\end{aligned}$$

and $\Phi^*_{0,1}(S,\bar{\theta}_{22})$ *is the* $(q_1 \times q_1)$ *left-upper bloc diagonal submatrix of the* $(q \times q)$ *matrix* $\Phi^*_0(S,\bar{\theta}_{22})$.

Proof. see Appendix A. □

Although calibrator's knowledge $\bar{\theta}_{22}$ is more often than not considered as free of parameter uncertainty thanks to the availability of a large bunch of empirical evidence, it is worth noticing that the above results remain valid even when it is explicitly acknowleged that in practice the exact value $\bar{\theta}_{22}$ is not known but only a root-T consistent estimator of it is available to define the simulator. More precisely, it can be easily shown that one can replace the value $\bar{\theta}_{22}$ of the nuisance parameters θ_{22} by a consistent estimator $\widehat{\theta}_{22,TS}$ such that $\sqrt{T}(\widehat{\theta}_{22,TS} - \bar{\theta}_{22}) = O_{P*}(1)$ without modifying the asymptotic probability distribution of the PII estimator. The detailled proof is not provided here since this result can actually be understood as a corollary of a very general adaptivity principle for extremum estimation. While the partial encompassing property precisely means that consistency of the estimators of the parameters of interest is not impaired by the chosen value for nuisance parameters, it also means by the same token that asymptotic probability distributions of these estimators do not depend on a specific choice of a root-T consistent estimator of the nuisance parameters.

3.3. Identifying the moments to match

The general idea is to start from a set of moments to match that are suggested by economic theory or any other features of the data the econometrician wishes to reproduce. The key hypothesis to test is the full encompassing property of theses moments by the structural model of interest. When the full encompassing hypothesis is rejected, one has to find a well suited selection, or more generally projection, of the initial instrumental characteristics in order to get at least partial encompassing. As explained above, this is the condition required to build a consistent partial indirect estimator as well as reliable predictions under hypothetical policy interventions.

Proposition 3.6. *Under assumptions* (A1)–(A9) *and the null hypothesis* H_0 *of full-encompassing of* $\mathcal{N}_\beta$ *by* (B1)

$$\xi_{T,S} = T \min_{\theta \in \Theta} \left[\widehat{\beta}_T - \frac{1}{S} \sum_{s=1}^{S} \widetilde{\beta}^s_T(\theta_1, \theta_2) \right]' \widehat{\Omega}^*_T(S) \left[\widehat{\beta}_T - \frac{1}{S} \sum_{s=1}^{S} \widetilde{\beta}^s_T(\theta_1, \theta_2) \right],$$

where $\widehat{\Omega}^*_T(S)$ *is a consistent estimator of the optimal metric* $\Omega^*(S) = \Phi^*_0(S)^{-1}$ *defined in Proposition* 3.4, *is asymptotically distributed as a* χ^2 *with* $(q-p)$ *degrees of freedom where* $q = \dim \beta$ *and* $p = \dim \theta$.

Proof. see Appendix A. □

The associated specification test of asymptotic level α is defined by the critical region:

$$\mathcal{W}_a = \{\xi_{T,S} > \chi^2_{1-\alpha}(q-p)\}.$$

In case of rejection, we may look for a reduction through an appropriate projection of the set of moments. This is based on the following partial encompassing test.

Proposition 3.7. *Under assumptions* (A1)–(A4), (A10)–(A14) *and the null hypothesis* $H_0(\bar{\theta}_{22})$ *of partial encompassing of* $\mathcal{N}_\beta$ *by* (B1)

$$\xi^1_{T,S}(\bar{\theta}_{22}) = T \min_{\theta_1,\theta_{21}\in\Theta_1\times\Theta_{21}} \left[\widehat{\beta}_{1,T} - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}^s_{1,T}(\theta_1,\theta_{21},(\bar{\theta}_{22}))\right]' \widehat{\Omega}^*_{1,T}(S)$$
$$\times\left[\widehat{\beta}_{1,T} - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}^s_{1,T}(\theta_1,\theta_{21},(\bar{\theta}_{22}))\right],$$

where $\Omega^*_{1,T}(S)$ *is a consistent estimator of the optimal metric* $\Omega^*_1(S,\bar{\theta}_{22}) = \Phi^*_{0,1}(S,\bar{\theta}_{22})^{-1}$ *defined in Proposition* 3.5, *is asymptotically distributed as a* χ^2 *with* $(q_1 - p_1 - p_{21})$ *degrees if freedom where* $q_1 = \dim\beta_1, p_1 = \dim\theta, p_{21} = \dim\theta_{21}$.

The proof is omitted here since it is a simple extension of the previous one. The associated specification test of asymptotic level α is defined by the following critical region:

$$\mathcal{W}^1_a = \{\xi^1_{T,S}(\bar{\theta}_{22}) > \chi^2_{1-\alpha}(q_1 - p_1 - p_{21})\}.$$

The previous result is not modified if θ_{22} is replaced by a consistent estimator $\widehat{\theta}_{22,TS}$ such that $\sqrt{T}(\widehat{\theta}_{22,TS} - \bar{\theta}_{22}) = \mathrm{O}_{\mathrm{P}*}(1)$. In case of rejection of any trial run of partial encompassing, the pair (structure model, instrumental model) is inadequate and has to be changed. However, it may also be the case that several pairs lead to acceptation.

4. Sequential partial indirect inference

The previous sections have shown how a well-driven PII estimation strategy may yield a consistent estimator for the structural parameters of interest θ_1 given the nuisance parameters θ_2, partly estimated (θ_{21}) and partly calibrated (θ_{22}). With this estimator in hands, one can now evaluate the model through some additional dimensions of interest. These additional dimensions are summarized by an instrumental model N_ψ, the parameters of which are ψ as characterized by an extremum estimation defined as a minimizer of a criterion $M_T(\underline{y}_T, \underline{x}_T, \psi)$. Typically, this criterion is the loss function used to assess in a second step of verification how true the structural model is with respect to the stylized facts of interest, under the maintained assumption of consistent first step estimation of the parameters of interest.

In the case of moment conditions, the corresponding criterion is defined as:

$$M_T(\underline{y}_T, \underline{x}_T, \psi) = \frac{1}{2}\left(\frac{1}{T}\sum_{t=1}^{T} k(\underline{y}_t, \underline{x}_t) - \psi\right)'\left(\frac{1}{T}\sum_{t=1}^{T} k(\underline{y}_t, \underline{x}_t) - \psi\right),$$

where $k(\cdot)$ are the moments of interest. The estimators $\widehat{\psi}_T$ and $\widetilde{\psi}_{TS}(\theta_1,\theta_2)$ associated with the criterion are:

$$\widehat{\psi}_T = \arg\min_{\psi\in\Psi} M_T(\underline{y}_T, \underline{x}_T, \psi),$$

$$\widetilde{\psi}^s_T(\theta_1,\theta_2) = \arg\min_{\psi\in\Psi} M_T(\underline{\widetilde{y}}^s_T(\theta, z^s_0), \underline{x}_T, \psi),$$

$$\widetilde{\psi}_{TS}(\theta_1,\theta_2)=\frac{1}{S}\sum_{s=1}^{S}\widetilde{\psi}_T^s(\theta_1,\theta_2).$$

Under usual regularity conditions, these estimators converge uniformly in (θ_1,θ_2) to:

$$P_0\lim_{T\to\infty}\widehat{\psi}_T=\psi^0=\psi(P_0),$$

$$P_*\lim_{T\to\infty}\widetilde{\psi}_{TS}=\widetilde{\psi}(\theta_1,\theta_2).$$

An evaluation of the structural model can then be performed by measuring a distance between the empirical instrumental parameters $\widehat{\psi}$ and the theoretical one $\widetilde{\psi}(\theta_1,\theta_2)$. In the case of full-encompassing, this corresponds to a Wald test and the statistic test is given by

$$T(\widehat{\psi}_T-\widetilde{\psi}^k_{TS}(\widehat{\theta}_{1,TS},\widehat{\theta}_{2,TS}))'\widehat{\Omega}^{\psi}_T(\widehat{\psi}_T-\widetilde{\psi}_{TS}(\widehat{\theta}_{1,TS},\widehat{\theta}_{2,TS})),$$

where $\widehat{\Omega}^{\psi}_T$ is a estimator of $\Omega^{\psi,*}$ and

$$\Omega^{\psi_*}=\Phi_0^{\psi,*}(S)^{-1},$$

$$\Phi_0^{\psi,*}(S)=[A,I]\Phi_0^*(S)[A,I]',$$

$$A=\left[-\frac{\partial\widetilde{\psi}}{\partial\theta'}(\theta_1^0,\bar{\theta}_2)'(W(S,\Omega^*))^{-1}\frac{\partial\widetilde{\beta}'}{\partial\theta}(\theta_1^0,\bar{\theta}_2)\Omega^*\right]$$

with $\Phi_0^*(S)$ defined in Appendix B.

This statistic is asymptotically distributed as a χ^2 with dim (ψ) degrees of freedom. Our proposed approach is then a two steps procedure. At the first step, we estimate the parameters of interest and at the second step, we evaluate the structural model. For this reason, we call this procedure as SPII. This typically corresponds to the Christiano and Eichenbaum (1992) strategy when they come to the second step of model verification about "labor-market moments". Note also that this two-step kind of approach is similar in spirit to what is done in Schorfheide (2000) when several loss functions are proposed to assess the discrepancy between DSGE model predictions and an overall posterior distribution of population characteristics that the researcher is trying to match.

Let us now consider the case of partial encompassing. As discussed above with the DSGE illustration, an estimator of the nuisance parameter vector θ_{22} can be obtained through the instrumental model of interest $\mathcal{N}_{\psi}$. We then define the estimator $\widehat{\theta}_{22,TS}$ as follows:

$$\widehat{\theta}_{22,TS}=\arg\min_{\theta_{22}\in\Theta_{22}}(\widehat{\psi}_T-\widetilde{\psi}_{TS}(\widehat{\theta}^1_{TS}(\bar{\theta}_{22}),\theta_{22}))'\widehat{\Omega}^{\psi}(\widehat{\psi}_T-\widetilde{\psi}_{TS}(\widehat{\theta}^1_{TS}(\bar{\theta}_{22}),\theta_{22})),$$

for a given initial value $\bar{\theta}_{22}$. Note that the estimator $\widehat{\theta}^1_{TS}(\bar{\theta}_{22})$ of (θ_1,θ_{21}) that we use here is the one associated to the initially calibrated value $\bar{\theta}_{22}$ and not to the fitted value θ_{22} produced at the verification stage. Once more, the idea is to disentangle the calibration and the verification steps. Of course, nothing would prevent us to make an eventual comparison of the competing empirical assessments of parameters θ_{22}. However, this should not be the focus of our interest since they are typically nuisance parameters.

Following Newey (1984), we can show the following proposition with assumptions (A.15)–(A.20) stated in the appendix:

Proposition 4.1. *Under the null hypothesis of the instrumental model $\mathcal{N}_\psi$ and assumptions* (A1)–(A4), (A15)–(A20), *the optimal II estimator $\widehat{\theta}^*_{22,TS}$ is obtained with the weighting matrix defined below. It is asymptotically normal, when S is fixed and T goes at infinity*:

$$\sqrt{T}(\widehat{\theta}^*_{22,TS} - \theta^*_{22}) \xrightarrow{\text{D}} \mathcal{N}(0, W^\psi(S, \Omega^{\psi,*})),$$

where

$$\theta^*_{22} = \underset{T\to\infty}{P_0} \lim \widehat{\theta}^*_{22,TS},$$

$$W^\psi(S, \Omega^{\psi,*}) = \left\{ \frac{\partial \widetilde{\psi}'}{\partial \theta_{22}}(\theta^0_1, \bar{\theta}_{21}, \theta^*_{22})(\Phi^{\psi,*}_0(S))^{-1} \frac{\partial \widetilde{\psi}}{\partial \theta'_{22}}(\theta^0_1, \bar{\theta}_{21}, \theta^*_{22}) \right\}^{-1},$$

$$\Omega^{\psi,*} = \Phi^{\psi,*}_0(S)^{-1},$$

$$\Phi^{\psi,*}_0(S) = [A, I]\Phi^*_0[A, I]',$$

$$A = \left[-\frac{\partial \widetilde{\psi}}{\partial \begin{pmatrix} \theta_1 \\ \bar{\theta}_{21} \end{pmatrix}'}(\theta^0_1, \bar{\theta}_{21}, \theta^*_{22})(W^*_{1,S}(\bar{\theta}_{22}))^{-1} \frac{\partial \widetilde{\beta}'_1}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}}(\theta^0_1, \bar{\theta}_{21}, \theta^*_{22})\Omega^*_1(\bar{\theta}_{22}) \right]$$

and Φ^*_0 is defined in Appendix B.

Proof. Appendix B. □

It should be emphasized that the asymptotic distribution given by Proposition 4.1 holds only for the same simulated values ε^s_t, $t = 1, \ldots, T$, $s = 1, \ldots, S$ for both instrumental models $\mathcal{N}_\beta$ and $\mathcal{N}_\psi$.

The SPII procedure is then the following for the partial encompassing case. At the first step, the estimators of $\theta^0_1(\bar{\theta}_{22})$ and $\bar{\theta}_{21}(\bar{\theta}_{22})$ are given by minimizing the following objective function:

$$J_{1,TS}(\theta_1(\bar{\theta}_{22}), \theta_{21}(\bar{\theta}_{22})) = [\widehat{\beta}_{1,T} - \widetilde{\beta}_{1,TS}(\theta_1, \theta_{21}, \bar{\theta}_{22})]' \times \widehat{\Omega}_{1,T}[\widehat{\beta}_{1,T} - \widetilde{\beta}_{1,TS}(\theta_1, \theta_{21}, \bar{\theta}_{22})]$$

for a given $\bar{\theta}_{22}$.

At the second step, the estimator of the nuisance parameters θ_{22} is given by minimizing the following objective function:

$$J_{2,TS}(\theta_{22}) = (\widehat{\psi}_T - \widetilde{\psi}_{TS}(\widehat{\theta}^1_{TS}(\bar{\theta}_{22}), \theta_{22}))'\widehat{\Omega}^\psi(\widehat{\psi}_T - \widetilde{\psi}_{TS}(\widehat{\theta}^1_{TS}(\bar{\theta}_{22}), \theta_{22})).$$

An evaluation of the structural model can then be performed by measuring a distance between the empirical instrumental parameters $\widehat{\psi}$ and the theoretical one $\widetilde{\psi}_{TS}(\theta_1, \theta_{21}, \theta_{22})$. In this context, the test corresponds to an overidentifying restrictions test. The test statistic is given by

$$TJ_{2,TS}(\widehat{\theta}_{22,TS}).$$

This statistic is asymptotically distributed as a χ^2 with $(\dim(\psi) - \dim(\theta_{22}))$ degrees of freedom.

5. Concluding remarks

The SPII methodology proposed in this paper aims at reconciling the calibration and verification steps proposed by the calibrationnist approach with their econometric counterparts, that is, estimation and testing procedures. We propose a general framework of multistep estimation and testing:

- First, for a given (calibrated) value $\bar{\theta}_{22}$ of some nuisance parameters, a consistent asymptotically normal estimator $\widehat{\theta}^{1}_{1,TS}(\bar{\theta}_{22})$ of the vector θ_1 of parameters of interest is obtained by partial indirect inference. A pseudo-true value $\bar{\theta}_{21}$ of some other nuisance parameters may also be consistently estimated by the same token.
- Second, overidentification of the vector (θ_1, θ_{21}) of structural parameters by the selected instrumental moments β_1 provides a specification test of the pair (structural model, instrumental model).
- Finally, the verification step, including a statistical assessment of the calibrated value $\bar{\theta}_{22}$, can be performed through another instrumental model $\mathcal{N}_{\psi}$.

The proposed formalization enables us to answer most of the common statistical criticisms about the calibration methodology, insofar as one succeeds to split the model in some true identifying moment conditions and some nominal assumptions. The main message is twofold. First, acknowledging that any structural model is misspecified while aiming at producing consistent estimators of the true unknown value of some parameters of interest as well as robust predictions, one should rely, as informally advocated in calibration exercises, on parsimonious and well chosen dimensions of interest. Second, in so doing, it may be the case that simultaneous joint estimation of the true unknown value of the parameters of interest as well as of the pseudo-true value of the nuisance parameters is impossible. In this context, one should resort to a two step procedure that we call sequential partial indirect inference (SPII). This basically introduces a general loss function. This again corresponds to a statistical formalization of the common practice in calibration exercises using previous estimates and a priori selection.

Acknowledgements

We thank Laurence Broze, Lars Peter Hansen and Frank Schorfheide for helpful comments. We thank two anonymous referees for useful comments which improved our paper. The second and the third authors would like to acknowledge the financial support from the Fonds Québécois de la Recherche sur la Société et la Culture (FQRSC).

Appendix A

Proof of Proposition 3.4. First-order conditions for the indirect estimator $\widehat{\theta}_{TS}$:

The first-order conditions corresponding to the optimization problem:

$$\min_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2} \left[\widehat{\beta}_T - \sum_{s=1}^{S} \widetilde{\beta}^{s}_T(\theta_1,\theta_2)\right]' \widehat{\Omega} \left[\widehat{\beta}_T - \sum_{s=1}^{S} \widetilde{\beta}^{s}_T(\theta_1,\theta_2)\right],$$

are:

$$\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widetilde{\beta}_T^{s\prime}}{\partial\theta}(\widehat{\theta}_{1,TS},\widehat{\theta}_{2,TS})\widehat{\Omega}_T\sqrt{T}\left[\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^{s}(\widehat{\theta}_{1,TS},\widehat{\theta}_{2,TS})\right]=0.$$

The expansion of the first-order conditions around the limit value $(\theta_1^{0\prime},\overline{\theta}_{21}^{\prime})'$ gives

$$\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widetilde{\beta}_T^{s\prime}}{\partial\theta}(\theta_1^0,\overline{\theta}_2)\Omega\sqrt{T}\left[\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^{s}(\theta_1^0,\overline{\theta}_2)-\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widetilde{\beta}_T^{s}}{\partial\theta'}(\theta_1^0,\bar{\theta}_2)\begin{pmatrix}\widehat{\theta}_{1,TS}-\theta_1^0\\ \widehat{\theta}_{2,TS}-\overline{\theta}_2\end{pmatrix}\right]$$
$$=\mathrm{o}_{\mathrm{P}_*}(1),$$

which leads to

$$\sqrt{T}\begin{pmatrix}\widehat{\theta}_{1,TS}-\theta_1^0\\ \widehat{\theta}_{2,TS}-\bar{\theta}_2\end{pmatrix}=\left\{\frac{\partial\widetilde{\beta}'}{\partial\theta}(\theta_1^0,\overline{\theta}_2)\Omega\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0,\overline{\theta}_2)\right\}^{-1}$$
$$\times\frac{\partial\widetilde{\beta}'}{\partial\theta}(\theta_1^0,\overline{\theta}_2)\Omega\sqrt{T}\left\{\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^{s}(\theta_1^0,\overline{\theta}_2)\right\}+\mathrm{o}_{\mathrm{P}_*}(1),$$

since under assumption (A9) we have

$$P_*\lim_{T\to+\infty}\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widetilde{\beta}_T^{s}}{\partial\theta}(\theta_1^0,\bar{\theta}_2)=\frac{\partial\widetilde{\beta}_T^{s\prime}}{\partial\theta}(\theta_1^0,\bar{\theta}_2)$$

and the transpose of the right-hand term is of full-column rank p.

Expansions of $\widehat{\beta}_T$ and $\widetilde{\beta}_T^{s}(\theta_1^0,\overline{\theta}_2)$:

We begin with the first-order conditions on the instrumental criterion:

$$\frac{\partial Q_T}{\partial\beta}(\underline{y}_T,\underline{x}_T,\widehat{\beta}_T)=0.$$

The expansion of the latter equation around the limit value β^0 yields:

$$\sqrt{T}\frac{\partial Q_T}{\partial\beta}(\underline{y}_t,\underline{x}_T,\beta^0)+\frac{\partial^2 Q_T}{\partial\beta\partial\beta'}(\underline{y}_T,\underline{x}_T,\beta^0)\sqrt{T}[\widehat{\beta}-\beta^0]=\mathrm{o}_{\mathrm{P}_0}(1),$$

which leads to

$$\sqrt{T}[\widehat{\beta}_T-\beta^0]=-J_0^{-1}\sqrt{T}\frac{\partial Q_T}{\partial\beta}(\underline{y}_T,\underline{x}_T,\beta^0)+\mathrm{o}_{\mathrm{P}_0}(1).$$

By using the same arguments

$$\sqrt{T}[\widetilde{\beta}_T^{s}(\theta_1^0,\overline{\theta}_2)-\widetilde{\beta}(\theta_1^0,\overline{\theta}_2)]=-J_0^{*-1}\sqrt{T}\frac{\partial Q_T}{\partial\beta}(\underline{\widetilde{y}}_T^{s}(\theta_1^0,\overline{\theta}_2,z_0^s),\underline{x}_T,\widetilde{\beta}(\theta_1^0,\overline{\theta}_2)+\mathrm{o}_{\mathrm{P}_*}(1)$$

and by full-encompassing hypothesis $H_0:\beta^0=\widetilde{\beta}(\theta_1,\overline{\theta}_2)$ we get

$$\sqrt{T}[\widetilde{\beta}_T^{s}(\theta_1^0,\overline{\theta}_2)-\beta^0]=-J_0^{*-1}\sqrt{T}\frac{\partial Q_T}{\partial\beta}(\underline{\widetilde{y}}_t^{s}(\theta_1^0,\overline{\theta}_2,z_0^s),\underline{x}_T,\beta^0)+\mathrm{o}_{\mathrm{P}_*}(1).$$

Asymptotic distribution of: $\sqrt{T}[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2)]$:

$$\sqrt{T}\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2)\right] = -J_0^{-1}\sqrt{T}\frac{\partial Q_T}{\partial \beta}(\underline{y}_T, \underline{x}_T, \beta^0) + J_0^{*-1}\sqrt{T}\frac{1}{S}\sum_{s=1}^{S}\frac{\partial Q_T}{\partial \beta}(\underline{\widetilde{y}}_T^s(\theta_1^0, \overline{\theta}_2, z_0^s), \underline{x}_T, \beta^0) + o_{P_*}(1).$$

Under assumption (A1)–(A9), $\sqrt{T}[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2)]$ is asymptotically normally distributed with mean zero and a covariance matrix given by $\Phi_0^*(S)$:

$$\Phi_0^*(S) = J_0^{-1}I_0J_0^{-1} + \frac{1}{S}J_0^{*-1}I_0J_0^{*-1} + \left(1 - \frac{1}{S}\right)J_0^{*-1}K_0^*J_0^{*-1} - J_0^{-1}K_0J_0^{*-1} - J_0^{*-1}K_0'J_0^{-1}.$$

As usual the optimal choice of the matrix Ω which minimizes the asymptotic variance of the II estimator is $\Omega^* = \Phi_0^*(S)^{-1}$ and the result of Proposition 3.4 follows. □

Proof of Proposition 3.7. The optimal value of the objective function is

$$\xi_{T,S} = T\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_t^s(\widehat{\theta}_{1,TS}, \widehat{\theta}_{2,TS})\right]'\widehat{\Omega}_T^*\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\widehat{\theta}_{1,TS}, \widehat{\theta}_{2,TS})\right],$$

where $(\widehat{\theta}_{1,TS}', \widehat{\theta}_{2,TS}')'$ corresponds to the optimal II estimator. The first-order expansion of $\xi_{T,S}$ around the limit value $(\theta_1^{0\prime}, \overline{\theta}_2')'$ gives

$$\xi_{T,S} = T\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2) - \frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0, \overline{\theta}_2)\begin{pmatrix}\widehat{\theta}_{1,TS} - \theta_1^0 \\ \widehat{\theta}_{2,TS} - \overline{\theta}_2\end{pmatrix}\right]'\Omega^* \times\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2) - \frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0, \overline{\theta}_2)\begin{pmatrix}\widehat{\theta}_{1,TS} - \theta_1^0 \\ \widehat{\theta}_{2,TS} - \overline{\theta}_2\end{pmatrix}\right] + o_{P_*}(1).$$

By using the asymptotic expansion of $\sqrt{T}\begin{pmatrix}\widehat{\theta}_{1,TS} - \theta_1^0 \\ \widehat{\theta}_{2,TS} - \overline{\theta}_2\end{pmatrix}$ around the limit value $(\theta_1^{0\prime}, \overline{\theta}_2')'$ previously given, we get

$$\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0, \overline{\theta}_2)\sqrt{T}\begin{pmatrix}\widehat{\theta}_{1,TS} - \theta_1^0 \\ \widehat{\theta}_{2,TS} - \overline{\theta}_2\end{pmatrix} = \frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0, \overline{\theta}_2)\left\{\frac{\partial\widetilde{\beta}'}{\partial\theta}(\theta_1^0, \overline{\theta}_2)\Omega^*\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0, \overline{\theta}_2)\right\}^{-1} \times\frac{\partial\widetilde{\beta}'}{\partial\theta}(\theta_1^0, \overline{\theta}_2)\Omega^*\sqrt{T}\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2)\right] + o_{P_*}(1)$$

and thus

$$\sqrt{T}\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2) - \frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0, \overline{\theta}_2)\begin{pmatrix}\widehat{\theta}_{1,TS} - \theta_1^0 \\ \widehat{\theta}_{2,TS} - \overline{\theta}_2\end{pmatrix}\right] = [Id_q - M]\sqrt{T}\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0, \overline{\theta}_2)\right] + o_{P_*}(1),$$

where M is the orthogonal projector on the space spanned by the columns of $\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0,\overline{\theta}_2)$ for the inner product Ω^* that is

$$M=\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0,\overline{\theta}_2)\left\{\frac{\partial\widetilde{\beta}'}{\partial\theta}(\theta_1^0,\overline{\theta}_2)\Omega^*\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0,\overline{\theta}_2)\right\}^{-1}\frac{\partial\widetilde{\beta}'}{\partial\theta}(\theta_1^0,\overline{\theta}_2)\Omega^*.$$

With these notations, the statistic $\xi_{T,S}$ is equal to

$$\xi_{T,S}=T\left[\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0,\overline{\theta}_2)\right]'\left(Id_q-M\right)'\Omega^*\left(Id_q-M\right)\left[\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0,\overline{\theta}_2)\right]+o_{P_*}(1).$$

As previously seen $\sqrt{T}[\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0,\overline{\theta}_2)]\overset{D}{\rightarrow}N[0,\Omega^{*-1}]$ as $T\rightarrow\infty$ and $\frac{\partial\widetilde{\beta}}{\partial\theta'}(\theta_1^0,\overline{\theta}_2)$ is full-column rank (p) which implies that

$$T\left[\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0,\overline{\theta}_2)\right]'(Id_q-M)'\Omega^*(Id_q-M)\left[\widehat{\beta}_T-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_t^s(\theta_1^0,\overline{\theta}_2)\right]\overset{D}{\rightarrow}\chi^2(q-p),$$

as $T\rightarrow\infty$ and the result of Proposition 3.7 follows. □

Proof of Proposition 3.5. First-order conditions for the indirect estimator $\widehat{\theta}^1_{1,TS}(\overline{\theta}_{22})$: The first-order conditions corresponding to the optimization problem:

$$\min_{(\theta_1,\theta_{21})\in\Theta_1\times\Theta_{21}}\left[\widehat{\beta}_{1,T}-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_{1,T}^s(\theta_1,\theta_{21},\widehat{\theta}_{22,TS})\right]'\widehat{\Omega}_{1,T}\left[\widehat{\beta}_{1,T}-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_{1,T}^s(\theta_1,\theta_{21},\widehat{\theta}_{22,TS})\right],$$

where $\widehat{\theta}_{22,TS}$ is a consistent estimator of the value $\overline{\theta}_{22}$ of the nuisance parameters θ_{22} and such that $\sqrt{T}(\widehat{\theta}_{22,TS}-\overline{\theta}_{22})=O_{P_*}(1)$, are

$$\begin{aligned}&\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widetilde{\beta}_{1,T}^{s'}}{\partial\binom{\theta_1}{\theta_{21}}}(\widehat{\theta}^s_{1,TS}(\bar{\theta}_{22}),\widehat{\theta}_{21,TS}(\bar{\theta}_{22}),\widehat{\theta}_{22,TS})\widehat{\Omega}_{1,T}\sqrt{T}\\&\quad\times\left[\widehat{\beta}_{1,T}-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_{1,T}^s(\widehat{\theta}^1_{1,TS}(\bar{\theta}_{22}),\widehat{\theta}^1_{21,TS}(\bar{\theta}_{22}),\widehat{\theta}_{22,TS})\right]=0.\end{aligned}$$

The expansion of the first-order conditions around the limit value $(\theta_1^{0'},\bar{\theta}_2')'$ gives

$$\begin{aligned}&\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widetilde{\beta}_{1,T}^{s'}}{\partial\binom{\theta_1}{\theta_{21}}}(\theta_1^0,\bar{\theta}_2)\Omega_1\sqrt{T}\left[\widehat{\beta}_{1,T}-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_{1,T}^s(\theta_1^0,\bar{\theta}_2)-\frac{1}{S}\sum_{s=1}^{S}\frac{\partial\widetilde{\beta}_{1,T}^s}{\partial\binom{\theta_1}{\theta_{21}}'}(\theta_1^0,\bar{\theta}_2)\begin{pmatrix}\widehat{\theta}^1_{1,TS}(\overline{\theta}_{22})-\theta_1^0\\ \widehat{\theta}^1_{21,TS}(\overline{\theta}_{22})-\overline{\theta}_{21}\end{pmatrix}\right]\\&\quad=o_{P_*}(1),\end{aligned}$$

since under $H_0^1 : \frac{\partial \widetilde{\beta}_1}{\partial \theta'_{22}}(\theta_1, \theta_2) = 0$. This leads to

$$\sqrt{T}\begin{bmatrix} \widehat{\theta}^1_{1,TS}(\bar{\theta}_{22}) - \theta^0_1 \\ \widehat{\theta}^1_{21,TS}(\bar{\theta}_{22}) - \bar{\theta}_{21} \end{bmatrix} = \left[\frac{\partial \widetilde{\beta}'_1}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}}(\theta^0_1, \bar{\theta}_{21})\Omega_1 \frac{\partial \widetilde{\beta}_1}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}'}(\theta^0_1, \bar{\theta}_{21})\right]^{-1}$$

$$\times \frac{\partial \widetilde{\beta}'_1}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}}(\theta^0_1, \bar{\theta}_{21})\Omega_1 \sqrt{T}\left[\widehat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S} \widetilde{\beta}^s_{1,T}(\theta^0_1, \bar{\theta}_2)\right] + \mathrm{o}_{\mathrm{P}_*}(1),$$

since under assumption (A14), we have

$$P_* \lim_{T \to +\infty} \frac{\partial \widetilde{\beta}^s_{1,T'}}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}}(\theta^0_1, \bar{\theta}_2) = \frac{\partial \widetilde{\beta}'_1}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}}(\theta^0_1, \bar{\theta}_{21})$$

and under H_0^1:

$$\frac{\partial \widetilde{\beta}_1}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}'}(\theta^0_1, \bar{\theta}_2) = \frac{\partial \widetilde{\beta}_1}{\partial \begin{pmatrix} \theta_1 \\ \theta_{21} \end{pmatrix}'}(\theta^0_1, \bar{\theta}_{21}),$$

is of full-column rank $(p_1 + p_{21})$.

Expansions of $\widehat{\beta}_{1,T}$ and $\widetilde{\beta}^s_{1,T}(\theta^0_1, \bar{\theta}_2)$:

We begin with the expansion of the first-order conditions on the instrumental model around the limit value β^0:

$$\frac{\partial Q_T}{\partial \beta}(\underline{y}_t, \underline{x}_T, \widehat{\beta}_T) = 0.$$

The expansion of the latter equation around the limit value β^0 yields

$$\sqrt{T}\frac{\partial Q_T}{\partial \beta}(\underline{y}, \underline{x}, \beta^0) + \frac{\partial^2 Q_T}{\partial \beta \partial \beta'}(\underline{y}, \underline{x}_T, \beta^0)\sqrt{T}[\widehat{\beta}_T - \beta^0] = \mathrm{o}_{\mathrm{P}_0}(1),$$

which leads to

$$\sqrt{T}[\widehat{\beta}_T - \beta^0] = -J_0^{-1}\sqrt{T}\frac{\partial Q_T}{\partial \beta}(\underline{y}, \underline{x}_t, \beta^0) + \mathrm{o}_{\mathrm{P}_0}(1).$$

By using the same argument:

$$\sqrt{T}[\widetilde{\beta}^s_T(\theta^0_1, \bar{\theta}_2) - \widetilde{\beta}^0(\bar{\theta}_{22})] = -J_0^{*-1}(\bar{\theta}_{22})\sqrt{T}\frac{\partial Q_T}{\partial \beta}(\underline{\widetilde{y}}^s_t(\theta^0_1, \bar{\theta}_2, z^s_0), \underline{x}_T, \widetilde{\beta}^0(\bar{\theta}_{22})) + \mathrm{o}_{\mathrm{P}_*}(1).$$

Asymptotic distribution of $\sqrt{T}[\widehat{\beta}_{1,T} - \frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}^s_{1,T}(\theta^0_1, \bar{\theta}_2)]$:

We have

$$\sqrt{T}\left[\widehat{\beta}_T-\beta^0-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1,\bar{\theta}_2)+\widetilde{\beta}^0(\bar{\theta}_{22})\right]$$
$$=-J_0^{-1}\sqrt{T}\frac{\partial Q_T}{\partial\beta}(\underline{y}_t,\underline{x}_T,\beta^0)+J_0^{*-1}(\bar{\theta}_{22})\sqrt{T}\frac{1}{S}\sum_{s=1}^{S}\frac{\partial Q_T}{\partial\beta}(\underline{\widetilde{y}}_t^s(\theta_1^0,\bar{\theta}_2,z_0^s),\underline{x}_T,\widetilde{\beta}^0(\bar{\theta}_{22}))+o_{P_*}(1).$$

The statistic $\sqrt{T}[\widehat{\beta}_T-\beta^0-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_T^s(\theta_1^0,\bar{\theta}_2)+\widetilde{\beta}^0(\bar{\theta}_{22})]$ is asymptotically normally distributed with mean zero and a covariance matrix given by $\Phi_0^*(S,\bar{\theta}_{22})$:

$$\Phi_0^*(S,\bar{\theta}_{22})=J_0^{-1}I_0J_0^{-1}+\frac{1}{S}J_0^{*-1}(\bar{\theta}_{22})I_0^*(\bar{\theta}_{22})J_0^{*-1}(\bar{\theta}_{22})+\left(1-\frac{1}{S}\right)J_0^{*-1}(\bar{\theta}_{22})K_0^*(\bar{\theta}_{22})J_0^{*-1}(\bar{\theta}_{22})$$
$$-J_0^{-1}K_0(\bar{\theta}_{22})J_0^{*-1}(\bar{\theta}_{22})-J_0^{*-1}(\bar{\theta}_{22})K_0'(\bar{\theta}_{22})J_0^{-1}.$$

Let $\Phi_{0,1}^*(S,\bar{\theta}_{22})$ be the $(q_1\times q_1)$ left-upper bloc diagonal sub-matrix of the $(q\times q)$ matrix $\Phi_0^*(S,\bar{\theta}_{22})$. By the partial-encompassing hypothesis H_0^1: $\beta_1^0=\widetilde{\beta}_1(\theta_1^0,\bar{\theta}_{21})$ we get

$$\sqrt{T}\left[\widehat{\beta}_{1,T}-\beta_1^0-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_{1,T}^s(\theta_1^0,\bar{\theta}_{21})+\widetilde{\beta}_1(\theta_1^0,\bar{\theta}_{21})\right]=\sqrt{T}\left[\widehat{\beta}_{1,T}-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_{1,T}^s(\theta_1^0,\bar{\theta}_{21})\right].$$

The statistic $\sqrt{T}[\widehat{\beta}_{1,T}-\frac{1}{S}\sum_{s=1}^{S}\widetilde{\beta}_{1,T}^s(\theta_1^0,\bar{\theta}_{21})]$ is asymptotically normally distributed with mean zero and a covariance matrix given by $\Phi_{0,1}^*(S,\bar{\theta}_{22})$. As usual the optimal choice of the matrix Ω_1 which minimizes the asymptotic covariance of the II estimator based on the sub-vector binding function is $\Omega_1^*(\bar{\theta}_{22})=\Phi_{0,1}(S,\bar{\theta}_{22})^{-1}$ and the result of Proposition 3.5 follows. □

Appendix B

We define the vector

$$\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{y}_T,\underline{x}_T,\gamma)=\left(\frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{y}_T,\underline{x}_T,\beta)',\frac{1}{\sqrt{T}}\frac{\partial M_T}{\partial\psi}(\underline{y}_T,\underline{x}_T,\psi)'\right)',$$

where $\gamma=(\beta',\psi')'$. The corresponding vector for the structural model is

$$\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{\widetilde{y}}_T^s(\theta_1^0,\bar{\theta}_2,z_0^s),\gamma)=\left(\frac{1}{\sqrt{T}}\frac{\partial Q_T}{\partial\beta}(\underline{\widetilde{y}}_T^s(\theta_1^0,\bar{\theta}_2,z_0^s),\beta)',\frac{1}{\sqrt{T}}\frac{\partial M_T}{\partial\psi}(\underline{\widetilde{y}}_T^s(\theta_1^0,\bar{\theta}_2,z_0^s),\psi)'\right)'.$$

We make the following assumptions:

(A15) $$\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{y}_T,\underline{x}_T,\gamma^0),$$

is asymptotically normally distributed with mean zero and with an asymptotic covariance matrix A_0.

(A16) $$B_0=P_0\lim_{T\to\infty}\frac{\partial^2V_T}{\partial\gamma\partial\gamma'}(\underline{y}_T,\underline{x}_T,\gamma^0),$$

(A17) $$\lim_{T\to+\infty} \mathrm{Cov}_*\left\{\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{y}_T,\underline{x}_T,\gamma^0),\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{\tilde{y}}^s_T(\theta^0_1,\bar{\theta}_2,z^s_0),\gamma^0)\right\}=C_0,$$

independent of the initial values $z^s_0, s=1,\ldots,S$.

(A18) $$\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{\tilde{y}}^s_T(\theta^0_1,\bar{\theta}_2,z^s_0),\gamma^0),$$

is asymptotically normally distributed with mean zero and with an asymptotic covariance matrix A^*_0 and independent of the initial values $z^s_0, s=1,\ldots,S$.

(A19) $$B^*_0 = P_*\lim_{T\to\infty}\frac{\partial^2 V_T}{\partial\gamma\partial\gamma'}(\underline{\tilde{y}}^s_T(\theta^0_1,\bar{\theta}_2,z^s_0),\gamma^0).$$

(A20) $$\lim_{T\to+\infty} \mathrm{Cov}_*\left\{\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{\tilde{y}}^s_T(\theta^0_1,\bar{\theta}_2,z^s_0),\gamma^0),\frac{1}{\sqrt{T}}\frac{\partial V_T}{\partial\gamma}(\underline{\tilde{y}}^s_T(\theta^0_1,\bar{\theta}_2,z^s_0),\gamma^0)\right\}=C^*_0,$$

independent of the initial values z^s_0 and z^ℓ_0, for $s\neq\ell$.

We can show that

$$\Phi^*_0(S) = B_0^{-1}A_0B_0^{-1}+\frac{1}{S}B_0^{*-1}A^*_0B_0^{*-1}+\left(1-\frac{1}{S}\right)B_0^{*-1}C^*_0B_0^{*-1}-B_0^{-1}C_0B_0^{*-1}-B_0^{*-1}C'_0B_0^{-1}.$$

References

Bansal, R., Gallant, R.A., Hussey, R., Tauchen, G., 1995. Nonparametric estimation of structural models for high-frequency currency market data. Journal of Econometrics 66, 251–287.

Beaudry, P., Guay, A., 1996. What do interest rates reveal about the functioning of real business cycle models? Journal of Economic Dynamics and Control 20, 1661–1682.

Bonomo, M., Garcia, R., 1994. Can a well-fitted asset-pricing model produce mean reversion? Journal of Applied Econometrics 9, 19–29.

Burnside, C., Eichenbaum, M., 1996. Factor-hoarding and the propagation of business-cycle shocks. American Economic Review 86, 1154–1174.

Canova, F., 1994. Statistical inference in calibrated models. Journal of Applied Econometrics 9, 123–144.

Cechetti, S.G., Lam, P., Mark, N.C., 1990. Evaluating empirical tests of asset pricing models: alternative interpretations. American Economic Review 80, 48–51.

Cechetti, S.G., Lam, P., Mark, N.C., 1993. The equity premium and the risk free rate: matching the moments. Journal of Monetary Economics 31, 21–45.

Christiano, L., 1988. Why does inventory investment fluctuate so much. Journal of Monetary Economics 21, 247–280.

Christiano, L., Eichenbaum, M., 1992. Current real business cycle theories and aggregate labor market fluctuations. American Economic Review 82, 430–450.

Dejong, D.N., Fisher, I.B., Whiteman, C.H., 1996. A Bayesian approach to calibration. Journal of Business and Economic Statistics 14, 1–9.

Dridi, R., Renault, E., 1998. Semiparametric indirect inference. Mimeo, Université de Toulouse.

Duffie, D., Singleton, K.J., 1993. Simulated moments estimation of Markov models of asset prices. Econometrica 61, 929–952.

Gallant, R.A., Tauchen, G., 1996. Which moments to match. Econometric Theory 12, 657–681.

Gallant, R.A., Hsieh, R., Tauchen, G., 1997. Estimating SV models with diagnostic. Journal of Econometrics 81, 159–192.

Geweke, J., 1999. Computational experiments and reality. Working Paper, University of Iowa.

Gouriéroux, C., Monfort, A., 1995. Testing, encompassing and simulating dynamic econometric models. Econometric Theory 11, 195–228.

Gouriéroux, C., Monfort, A., Trognon, A., 1984. Pseudo maximum likelihood methods: theory. Econometrica 52, 681–700.

Gouriéroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, 85–118.

Gouriéroux, C., Renault, E., Touzi, N., 2000. Calibration by simulation for small sample bias correction. In: Mariano, de.R., Schuermann, T., Weeks, M. (Eds.), Simulation-based Inference in Econometrics Methods and Applications. Cambridge University Press, Cambridge.

Gregory, A.W., Smith, G.W., 1990. Calibration as estimation. Econometric Reviews 9, 57–89.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hansen, G., 1985. Indivisible labor and the business cycle. Journal of Monetary Economics 16, 309–328.

Hansen, L.P., Heckman, J.J., 1996. The empirical foundations of calibration. Journal of Economic Perspectives 10, 87–104.

Hansen, L.P., Sargent, T.J., 1979. Formulating and estimating dynamic linear rational expectations models. Journal of Economic Dynamics and Control 2, 7–46.

Hansen, L.P., Sargent, T.J., 1993. Seasonality and approximation errors in rational expectations models. Journal of Econometrics 55, 21–55.

Hoover, K.D., 1995. Facts and artifacts: calibration and empirical assessment of real-business-cycle models. Oxford Economic Papers 47, 24–44.

Kim, K., Pagan, A., 1995. The econometrics of calibrated models. In: Pearan, M.H., Wickens, M. (Eds.), Handbook of Applied Econometrics, pp. 356–390.

King, R.G., Plosser, C.I., Rebelo, S., 1988a. Production, growth and business cycles: I. The basic neoclassical model. Journal of Monetary Economics 21, 195–232.

King, R.G., Plosser, C.I., Rebelo, S., 1988b. Production, growth and business cycles: II. The basic neoclassical model. Journal of Monetary Economics 21, 309–342.

King, R.G., Plosser, C.I., Stock, J.H., Watson, M., 1991. Stochastic trends and economic fluctuations. American Economic Review 81, 819–840.

Kydland, F.E., Prescott, E.C., 1982. Time to build and aggregate fluctuations. Econometrica 50, 1345–1370.

Kydland, F.E., Prescott, E.C., 1991. The econometrics of the general equilibrium approach to business cycles. Scandinavian Journal of Economics 93, 161–178.

Kydland, F.E., Prescott, E.C., 1996. The computational experiment: an econometric tool. Journal of Economic Perspectives 10, 69–86.

Lucas, R.E., 1978. Asset prices in an exchange economy. Econometrica 46, 1429–1445.

Lucas, R.E., 1980. Methods and problems in business cycle theory. Journal of Money, Credit and Banking 12, 696–715.

Lucas, R.E., 1987. Models of Business Cycles. Blackwell, Oxford.

Manuelli, R., Sargent, T.J., 1988. Models of business cycles: a review essay. Journal of Monetary Economics 22, 523–542.

Mehra, R., Prescott, E.C., 1985. The equity premium: a puzzle. Journal of Monetary Economics 15, 145–161.

Monfort, A., 1996. A reappraisal of misspecified econometric models. Econometric Theory 12, 597–619.

Newey, W.K., 1984. A method of moments interpretation of sequential estimators. Economic Letters 14, 201–206.

Prescott, E.C., 1983. Can the cycle be reconciled with a consistent theory of expectations or a progress report on business cycle theory. Working Paper no 239, Research Department, Federal Reserve Bank of Mineapolis.

Ruge-Murcia, F., 2003. Methods to estimate dynamic stochastic general equilibrium models. Working Paper, CIREQ, Université de Montréal.

Schorfheide, F., 2000. Loss function-based evaluation of DSGE models. Journal of Applied Econometrics 15, 645–670.

Tauchen, G., Zang, H., Liu, M., 1997. Volume, volatility and leverage: a dynamic analysis. Journal of Econometrics 74, 177–208.

Watson, M., 1993. Measure of fit for calibrated models. Journal of Political Economy 101, 1011–1041.

White, H., 1982. Maximum likelihood estimation of mis-specified models. Econometrica 50, 1–25.