Universität Augsburg

Institut für Volkswirtschaftslehre

Volkswirtschaftliche Diskussionsreihe

Business Cycle Uncertainty and Economic Welfare Revisited

Christopher Heiberger, Alfred Maußner

Beitrag Nr. 335, Juni 2018

Business Cycle Uncertainty and Economic Welfare Revisited

Christopher Heiberger[a] and Alfred Maußner[b]

[a]University of Augsburg, Department of Economics, Universitätsstraße 16, 86159 Augsburg, Germany, christopher.heiberger@wiwi.uni-augsburg.de

[b]Corresponding author, University of Augsburg, Department of Economics, Universitätsstraße 16, 86159 Augsburg, Germany, alfred.maussner@wiwi.uni-augsburg.de

Preliminary version.
Do not distribute without permission of the authors.
May 22, 2018

JEL classification: C63, D60, E32
Keywords:
Business cycles, Mean effect, Second order solution, Risk aversion, Welfare costs

Abstract

Cho, Cooley, and Kim (RED, 2015) (CCK) consider the welfare effects of removing multiplicative productivity shocks from real business cycle models. In a model that admits an analytical solution they argue convincingly that the positive welfare effect of removing uncertainty can be dominated by a negative mean effect arising from the optimal response of household labor supply. While the presentation of this model is quite elaborate, the details of their subsequent quantitative analysis of several versions of the standard real business cycle model remain vague. We lay out the general procedure of computing second-order accurate approximations of welfare gains or losses in the canonical dynamic stochastic general equilibrium model. In order to be able to consider mean preserving increases in the size of shocks we extend the computation of second-order approximations of the policy functions pioneered by Schmitt-Grohé and Uribe (JEDC, 2004). Our computations show that different from the results reported in CCK the mean effect never dominates the fluctuations effect. Welfare measures computed from weighted residuals methods confirm the logic behind our perturbation approach and verify the accuracy of our estimates.

1 Introduction

Would economic agents benefit if the business cycle could be removed? And if so, by how much? Lucas (1987) started this discussion by arguing these benefits are negligible. He considers a representative, risk-averse consumer facing a stochastic consumption stream. The consumer values this stream according to an additively separable intertemporal utility function with iso-elastic period utility. The expected mean of the stream grows at a constant rate and the fluctuations around this trend match the variance of trend deviations of quarterly real U.S. consumption. For plausible values of the coefficient of relative risk-aversion (his Table 2 considers values between 1 and 20) he estimates a welfare gain not exceeding 0.1 percent of annual consumption, i.e., about $ 8.5 in 1983.

Since then, many researchers have estimated the welfare costs of business cycles in more elaborate models.[1] These include departures from the specification of preferences and of the consumption process (as, e.g., Obstfeld (1994), Dolmas (1998), Tallarini (2000), and Barro (2009)), models with uninsurable idiosyncratic risk (as, e.g., İmrohoroğlu (1989), Krebs (2003), De Santis (2007), and Krusell et al. (2009)), the consideration of nominal frictions (as, e.g., Cho et al. (1997) and Galí et al. (2007)), endogenous growth (as, e.g., Barlevy (2004) and Heer and Maußner (2015)), and model uncertainty (Barillas et al. (2009)). The range of estimates is wide, from being close to Lucas' 0.1 percent to several orders of magnitude beyond. For instance, İmrohoroğlu (1989), p.1378 estimates 0.3 percent of average consumption, Krebs (2003), p. 862 finds 7.48 percent, Tallarini (2000), Table 3 calculates costs between 2.1 and 12.6 percent, and in the model of Barro (2009) the society would be willing to reduce GDP by about 20% to eliminate rare disasters.

In a recent paper Cho, Cooley, and Kim (2015) (henceforth CCK) contribute a methodological argument to the debate. They distinguish between the effect of uncertain consumption and leisure (their "fluctuations effect") and the effect of optimal adjustment of factor inputs to multiplicative shocks (their "mean effect"). The fluctuations effect captures risk-aversion. Risk-averse economic agents are characterized by concave utility functions, so that according to Jensen's inequality the expected utility of a lottery does not exceed the utility obtained from the expected outcome of the lottery. Therefore, they will benefit if the lottery is replaced by a certain stream of consumption and leisure equal to expected consumption and leisure from the lottery. Lucas (1987) only measures this effect. However, in a production economy, expected consumption and leisure are not exogenously given. Consider, for instance, the standard real business cycle model driven a by a single shock to total factor productivity (TFP). If economic agents respond optimally to this shock, the reduced form production function may become convex in the shock so that, again by Jensen's inequality, expected output exceeds production at the expected level of TFP. Therefore, removing uncertainty might be detrimental to economic welfare. CCK measure this mean effect by comparing life-time utility at the expected levels of consumption and leisure with those obtained in a deterministic environment.

The proper measurement of both the fluctuations and the mean effect requires that the means of the shocks that drive the economy are not unrelated to the standard deviations.

[1] See Barlevy (2005) for a survey of the literature.

To see this for the fluctuations effect consider the process $\{C_{t+s}\}_{s=0}^{\infty} := \{\exp(z_{t+s})\}_{s=0}^{\infty}$, where z_{t+s} is i.i.d. normally distributed with mean μ and standard deviation τ. Then $\mathbb{E}(C_{t+s}) = \exp(\mu + 0.5\tau^2)$ so that different degrees of uncertainty τ are associated with different amounts of expected consumption $\mathbb{E}(C_{t+s})$. Hence, in order not to mix level effects with the effect of removing uncertainty, one must assume $\mu(\tau) := -0.5\tau^2$. To see this for the mean effect, consider the example model from Section 2.2 of CCK where a representative agent chooses consumption c_t, labor supply n_t, and his future stock of capital k_{t+1} to maximize

$$\begin{aligned}
&U_t := \mathbb{E}_t \left\{ \sum_{s=0}^{\infty} \beta^s \frac{1}{1-\eta} (c_{t+s} - \alpha n_{t+s})^{1-\eta} \right\}, \quad \beta \in (0,1), \\
&\text{subject to} \\
&\left.\begin{aligned}
y_{t+s} &= A_{t+s} k_{t+s}^{\theta} n_{t+s}^{1-\theta}, \quad \theta \in (0,1), \\
k_{t+s+1} &= y_{t+s} - c_{t+s}, \\
k_t &> 0 \text{ given.}
\end{aligned}\right\} \text{ for } s = 1, 2, \dots
\end{aligned} \tag{E}$$

The optimal response of labor n_t to the TFP shock A_t yields the reduced form production function

$$y_t = \left(\frac{1-\theta}{\alpha} \right)^{\frac{1-\theta}{\theta}} A_t^{\frac{1}{\theta}} k_t$$

which is convex in TFP. Assume again that A_t is i.i.d. log-normal with mean μ and standard deviation τ. The standard assumption in the literature is $\mu = 0$, so that in the deterministic model with $\tau = 0$, TFP is equal to unity while expected TFP in the stochastic economy is equal to $\mathbb{E}(A_t) = \exp(0.5\tau^2) > 1$. Therefore, if we want to evaluate the mean effect we must eliminate the influence of τ on the expected level of TFP, requiring $\mu(\tau) := -0.5\tau^2$. More generally, thus, measuring the welfare effects of economic fluctuation necessitates models whose driving processes possess the mean preserving spread property, i.e., stochastic processes whose means depend on their standard deviation.

For the example given in (E) CCK derive analytically the restrictions which the parameters must meet for the overall effect of uncertainty to be negative. In more general models the question whether or not the mean effect dominates the fluctuations effect cannot be answered analytically. In Section 3 of their paper CCK employ numerical methods to compute the welfare effects of removing uncertainty in several well-known models: the stochastic growth model, the Hansen (1985) real business cycle model, and a simplified version of the two-country model of Backus et al. (1992). While their discussion of the example is elaborate and traceable, the description of their numerical method leaves two issues unsolved.

The first issue concerns the definition of the welfare measure. CCK refer to two papers by Stephanie Schmitt-Grohé and Martin Uribe (henceforth SGU)). Their first reference, Schmitt-Grohé and Uribe (2004b), and footnote 9 reveal that they employ the Matlab code of SGU and compute a second-order perturbation solution. Their second reference is Schmitt-Grohé and Uribe (2007). In this paper SGU compute *unconditional* welfare effects of various fiscal and monetary policy rules. They distinguish these from *conditional* effects which are based on the assumption that both the reference and the alternative economy start at the deterministic steady state.[2] CCK do not mention which measure they compute.

The second issue concerns the distribution of the shocks that drive the model. As required, CCK consider processes with the mean preserving spread property. Thus, the means of the innovations are generally different from zero and depend on the variance of the shocks. As a consequence, perturbing the variance also requires to adjust the mean appropriately. To the best of our knowledge, the Matlab programs of SGU assume means of zero which are independent of the perturbation parameter. CCK provide no discussion, whether or not they have adapted this code to fit their assumption, and their posted code does not compute their welfare measures.

The contribution of our paper, therefore, is first and foremost methodological. We extend the canonical stochastic general equilibrium (DSGE) model of Schmitt-Grohé and Uribe (2004b) to allow for stochastic processes that have the mean-preserving spread property. For this extended model we derive the second-order perturbation solution and provide Matlab code that implements this solution.[3] Our solution appropriately adjusts the level effect of the perturbation parameter. Thus, it also shifts the agent's value function required for welfare comparisons. Next we consider *conditional* and *unconditional* welfare measures that can be decomposed into the fluctuations and the mean effect and relate their computation to the perturbation solution of the canonical DSGE model. We apply our techniques to recompute the welfare effects of removing fluctuations from the model that underlies Figure 2 of CCK.[4] Different from this Figure, our results show that the mean effect never dominates the fluctuations effect. Finally, and in order to convince the reader of the logic behind our extension of the canonical DSGE model, we compare our perturbation solution of this model with those obtained from two different weighted residuals methods. Even though these methods produce much more accurate solutions in terms of Euler residuals, this increased degree of precision has negligible effects on the conditional welfare measure and on the unconditional measure for moderate degrees of uncertainty.

From here we proceed with a brief description of the benchmark real business cycle model in Section 2. This model, taken from CCK, serves as an example of the canonical DSGE model presented in Section 3.1 and as a framework to illustrate the computation of conditional and unconditional welfare measures in Section 3.2. In Section 3.3 we sketch the solution of this model via two weighted residuals methods. Section 4 presents our quantitative results. In particular, we provide conditional and unconditional welfare gains from removing fluctuations from the model of Section 2 and compare them to those of CKK. Section 5 concludes. The Appendix covers the detailed derivation of our second-order solution as well as the detailed presentation of our weighted residuals methods.

---

[2] Schmitt-Grohé and Uribe (2004a) is more elaborate about the computation of these measures. However, one has to study their Matlab code in order to resolve the details presented in Section 3.2.

[3] This code is also part of a toolbox that accompanies the textbook of Heer and Maußner (2009) and can be downloaded from `http://www.wiwi.uni-augsburg.de/vwl/maussner/dge_book_downloads/2nd_edition/CoRRAM-M_8March18_.zip`.

[4] CCK present their results in terms of graphs that display the welfare measure as a function of the coefficient of risk-aversion and of the variance of the TFP shock.

2 The model

In Section 3.1 of their paper CCK consider a standard real business cycle model similar to Hansen (1985). Output y_t is produced from capital k_t and labor n_t according to the production function

$$y_t = A_t k_t^{\theta} n_t^{1-\theta}. \tag{1}$$

The natural log of total factor productivity A_t is governed by an AR(1)-process

$$\ln A_t = \rho \ln A_{t-1} + \epsilon_t, \quad \epsilon_t \text{ i.i.d. } N\left(-\frac{\tau^2}{2(1+\rho)}, \tau^2\right). \tag{2}$$

The innovations in the process ϵ_t are i.i.d. normal. Its mean is chosen so that the unconditional expectation of A_t is equal to unity and independent of the standard deviation τ.[5] The current-period utility function u is of the Cobb-Douglas type and depends on consumption c_t and leisure n_t:[6]

$$u(c_t, 1-n_t) := \begin{cases} \frac{1}{1-\eta}\left[c_t^{\alpha}(1-n_t)^{1-\alpha}\right]^{1-\eta} & \text{for } \eta \neq 1, \\ \alpha \ln c_t + (1-\alpha)\ln(1-n_t) & \text{for } \eta = 1. \end{cases} \tag{3}$$

Thus, for given values of $\alpha \in (0,1)$ the coefficient of relative risk aversion,

$$CRRA := -\frac{\partial^2 u/\partial c^2}{\partial u/\partial c} c = 1 + \alpha(\eta - 1),$$

increases with η.

Capital depreciates at the constant rate $\delta \in (0,1]$, and the representative agent discounts future utilities at the rate $\beta \in (0,1)$. This agent chooses consumption, leisure, and the future stock of capital to maximize the expected life-time utility

$$V_t := \mathbb{E}_t \left\{ \sum_{s=0}^{\infty} \beta^s u(c_{t+s}, 1-n_{t+s}) \right\}, \quad \beta \in (0,1)$$

subject to the resource constraints

$$\left.\begin{aligned} y_{t+s} &\geq c_{t+s} + i_{t+s}, \\ k_{t+s+1} &= (1-\delta)k_{t+s} + i_{t+s}, \\ 0 &\leq c_{t+s}, k_{t+s+1}, \text{ and } n_{t+s} \in (0,1) \end{aligned}\right\} \text{ for } s = 0, 1, \ldots$$

and a given initial stock of capital k_t. Expectations $\mathbb{E}_t$ are conditioned on information available at time t. We provide the equilibrium conditions of this model in Appendix A.

Table 1 presents the parameter values chosen by CCK for those parameters which remain constant. CCK vary the values of η and τ in order to demonstrate their effect on economic welfare.

[5] See Cho et al. (2015) footnote 10.

[6] CCK use σ instead of η as a measure of risk aversion in the utility function. We employ the notation of Schmitt-Grohé and Uribe (2004b) for the canonical DSGE model. They designate the perturbation parameter by the Greek letter σ .

Table 1: Calibration of the model

Parameter	Description	Value
α	utility weight of consumption	0.35
β	discount factor	0.99
δ	rate of capital depreciation	0.025
ρ	autocorrelation of log of TFP shock	0.95

3 Second-order approximations of welfare effects in DSGE models

This section draws on the work of SGU. In Schmitt-Grohé and Uribe (2004b) they propose a canonical framework for dynamic stochastic general equilibrium (DSGE) models and develop formulas for a second-order approximate solution of this class of models. In Schmitt-Grohé and Uribe (2006, 2007) they develop measures for welfare comparisons.

3.1 Canonical DSGE model

Variables and equilibrium conditions. Let $\mathbf{x}_t \in \mathbb{R}^{n(x)}$ denote a vector of $n(x)$ endogenous state variables, i.e., variables whose period t value is given at the beginning of period t but evolves endogenously from period to period. The vector $\mathbf{y}_t \in \mathbb{R}^{n(y)}$ collects $n(y)$ endogenous variables being determined within period t. The driving forces of the model are $n(z)$ purely exogenous variables gathered in the vector $\mathbf{z}_t \in \mathbb{R}^{n(z)}$. The equilibrium conditions are given by

$$\mathbf{0}_{(n(x)+n(y))\times 1} = \mathbb{E}_t \mathbf{g}\left(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t\right), \tag{4a}$$

where $\mathbb{E}_t$ denotes mathematical expectations as of time t and $\mathbf{g} : \mathbb{R}^{2(n(x)+n(y)+n(z))} \to \mathbb{R}^{n(x)+n(y)}$. We denote the vector of all state variables by $\mathbf{w}_t := \begin{bmatrix}\mathbf{x}^T & \mathbf{z}^T\end{bmatrix}^T \in \mathbb{R}^{n(w)}$ where $n(w) = n(x) + n(z)$.

Driving process. Schmitt-Grohé and Uribe (2004b) specify the driving process of the exogenous variables as

$$\mathbf{z}_{t+1} = R\mathbf{z}_t + \sigma\Omega\boldsymbol{\nu}_{t+1}.$$

The matrix R has all its eigenvalues within the unit circle. The parameter $\sigma \geq 0$ is an arbitrary scalar factored out from the matrix Ω and plays the role of the perturbation parameter. Without loss of generality we will fix $\sigma = 1$ for the stochastic model in the following. The vector $\boldsymbol{\nu}_{t+1} \in \mathbb{R}^{n(z)}$ is distributed independently and identically with mean $\mathbf{0}_{n(z)\times 1}$ and variance $I_{n(z)}$,[7] and the matrix Ω determines the covariance of the innovations

$$\boldsymbol{\epsilon}_{t+1} := \sigma\Omega\boldsymbol{\nu}_{t+1}$$

[7] I_n denotes the identity matrix of dimension n.

since $\mathbb{E}_t(\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}_{t+1}^T) = \sigma^2\Omega\Omega^T$. For $\sigma = 0$ the process becomes deterministic and approaches the zero vector: $\lim_{t\to\infty} \mathbf{z}_t = \mathbf{0}_{n(z)\times 1}$. If the deterministic model has a stable solution at the point $\begin{bmatrix}\mathbf{x}^T & \mathbf{0}_{1\times n(z)} & \mathbf{y}^T\end{bmatrix}^T$ solving

$$\mathbf{0}_{(n(x)+n(y))\times 1} = \mathbf{g}\left(\mathbf{x}, \mathbf{0}_{n(z)\times 1}, \mathbf{y}, \mathbf{x}, \mathbf{0}_{n(z)\times 1}, \mathbf{y}\right),$$

one can invoke the implicit function theorem to approximate the solution for a nearby [8] stochastic model $\sigma = 1$.

To motivate our extension let us revert for the moment to the one-shock model in Section 2, where $\mathbf{z}_t = [\ln(A_t)]$. Note, if we assume $\epsilon_{t+1} \sim$ i.i.d. $N(-\tau^2/(2(1+\rho)), \tau^2)$, as in (2), an equivalent formulation is given by

$$\epsilon_{t+1} := -\frac{\tau^2}{2(1+\rho)} + \tau\nu_{t+1} = \tilde{\mu}(\tau) + \tau\nu_{t+1}, \;\; \nu_{t+1} \sim \text{ iid}N(0,1),$$

where $\tilde{\mu}(\tau) = -\frac{\tau^2}{2(1+\rho)}$. Setting further $R := \rho$ and $\Omega := \tau$ as well as $\sigma = 1$ for the stochastic model, the AR(1) specification (2) for log productivity implies that

$$\mathbf{z}_{t+1} = R\mathbf{z}_t + \tilde{\mu}(\sigma\Omega) + \sigma\Omega\boldsymbol{\nu}_{t+1}. \tag{4b}$$

Most importantly, note already here that this specification guarantees not only that the mean preserving spread property is met for the stochastic model with $\sigma = 1$, but also that it remains throughout valid when perturbing and taking derivatives with respect to σ.

Now, if $\tilde{\mu}$ was non-zero but independent of $\sigma\Omega$, one could simply convert the previous system to

$$\bar{\mathbf{z}}_{t+} = R\bar{\mathbf{z}}_t + \sigma\Omega\nu_{t+1} \text{ where } \bar{\mathbf{z}}_t := \mathbf{z}_t - \left(I_{n(z)} - R\right)^{-1}\tilde{\mu},$$

being equivalent to the process of SGU. Since the required transformation would then be independent of $\sigma\Omega$ and therefore in particular of the perturbation parameter σ, the second-order approximation derived in Schmitt-Grohé and Uribe (2004b) remains valid.

However, this is different if $\tilde{\mu}(\sigma\Omega)$ is non-constant as it is the case in the problem at hand. In order to adequately account for the effect of uncertainty in the model, i.e. in order to adequately cancel out any shift of the mean in productivity in the perturbation approach, the variance-covariance matrix and the mean vector in (4b) have to be perturbed simultaneously with σ.

Consequently, the second-order Taylor approximation for the stochastic model constructed around $\sigma = 0$ differs from the standard results in Schmitt-Grohé and Uribe (2004b), if the gradient or Hessian matrix of $\mu(\sigma) := \tilde{\mu}(\sigma\Omega)$ at $\sigma = 0$ are non-trivial. Obviously, for the present case where $\mu(\sigma) = -\frac{\sigma^2\tau^2}{2(1+\rho)}$ we get

$$\mu'(0) = -\frac{\tau^2\sigma}{1+\rho}\bigg|_{\sigma=0} = 0,$$

$$\mu''(0) = -\frac{\tau^2}{(1+\rho)} < 0.$$

[8] I.e. if the elements on the diagonal of $\Omega\Omega^T$ are not too large.

This motivates our extension of the shock process as part of the canonical DSGE model:

$$\mathbf{0}_{n(z)\times 1} = \mathbf{z}_{t+1} - R\mathbf{z}_t - \boldsymbol{\mu}(\sigma) - \sigma\Omega\boldsymbol{\nu}_{t+1}, \quad \boldsymbol{\nu}_{t+1} \text{ i.i.d. } N\left(\mathbf{0}_{n(z)\times 1}, I_{n(z)}\right), \tag{4c}$$

$$\boldsymbol{\mu}(\sigma = 0) = \mathbf{0}_{n(z)\times 1}, \tag{4d}$$

$$\boldsymbol{\mu}_\sigma(\sigma = 0) = \mathbf{0}_{n(z)\times 1}, \tag{4e}$$

$$\boldsymbol{\mu}_{\sigma\sigma}(\sigma = 0) \neq \mathbf{0}_{n(z)\times n(z)}. \tag{4f}$$

Approximate solution. The stationary point of the deterministic model is found as solution of

$$\mathbf{0}_{(n(x)+n(y))\times 1} = \mathbf{g}\left(\mathbf{x}, \mathbf{0}_{n(z)\times 1}, \mathbf{y}, \mathbf{x}, \mathbf{0}_{n(z)\times 1}, \mathbf{y}\right).$$

The solution of the perturbed model are vector valued functions $\mathbf{h}^x : \mathbb{R}^{n(x)+n(z)+1} \to \mathbb{R}^{n(x)}$ and $\mathbf{h}^y : \mathbb{R}^{n(x)+n(z)+1} \to \mathbb{R}^{n(y)}$ given by

$$\mathbf{x}_{t+1} = \mathbf{h}^x(\mathbf{w}_t, \sigma), \tag{5a}$$

$$\mathbf{y}_t = \mathbf{h}^y(\mathbf{w}_t, \sigma). \tag{5b}$$

Schmitt-Grohé and Uribe (2004b) show how to compute a second-order approximation of these functions at the point of expansion $\left[\mathbf{x}^T, \mathbf{0}_{1\times n(z)}, 0\right]^T$, i.e. at the stationary point of the deterministic model. The form of this solution reads

$$\mathbf{h}^x(\mathbf{x}_t, \mathbf{z}_t, \sigma) = \mathbf{x} + H^x_w(\mathbf{w}_t - \mathbf{w}) + \tfrac{1}{2}\left(I_{n(x)} \otimes (\mathbf{w}_t - \mathbf{w})^T\right) H^x_{ww}(\mathbf{w}_t - \mathbf{w}) + \tfrac{1}{2} H^x_{\sigma\sigma}\sigma^2, \tag{6a}$$

$$\mathbf{h}^y(\mathbf{x}_t, \mathbf{z}_t, \sigma) = \mathbf{y} + H^y_w(\mathbf{w}_t - \mathbf{w}) + \tfrac{1}{2}\left(I_{n(y)} \otimes (\mathbf{w}_t - \mathbf{w})^T\right) H^y_{ww}(\mathbf{w}_t - \mathbf{w}) + \tfrac{1}{2} H^y_{\sigma\sigma}\sigma^2. \tag{6b}$$

Setting $\sigma = 1$ in these equations delivers the solution of the stochastic model.

The Matlab code provided by Schmitt-Grohé and Uribe (2004b) for the computation of the matrices H^x_{ww}, H^y_{ww}, $H^x_{\sigma\sigma}$, and $H^y_{\sigma\sigma}$ rests on tensor formulas. In the Appendix B we apply a chain rule for the second derivative of a vector-valued composite function as proposed by Gomme and Klein (2011) and extend the computation of $H^x_{\sigma\sigma}$ and $H^y_{\sigma\sigma}$ to the case of perturbed means as specified in (4c). As can be expected, it turns out that $H^x_{\sigma\sigma}$ and $H^y_{\sigma\sigma}$ differ from the standard case by a term involving the Hessian of $\mu(\sigma)$, which we interpret as the effect from considering mean preserving spreads. This term is essential in order to adequately cancel out any effects in the policy functions from unwanted shifts in the means when alternating the degree of uncertainty in the model.

3.2 Welfare measures

Conditional measure. Suppose we have two different equilibrium time paths $i \in \{a, r\}$ of the model of Section 2. Let

$$V_{it} := \mathbb{E}_t\left\{\sum_{s=0}^{\infty} \beta^s \frac{1}{1-\eta} c_{it+s}^{\alpha(1-\eta)} (1 - n_{it+s})^{(1-\alpha)(1-\eta)}\right\} \tag{7}$$

denote the associated life-time utility, and note that it is conditioned on the given initial capital stock k_t and the known initial realization of the TFP shock $\ln A_t$. Let $i=r$ denote our reference equilibrium to which we want to compare an alternative equilibrium $i=a$. Schmitt-Grohé and Uribe (2004a), p. 17f define the welfare measure as the fraction of consumption which the representative agent would be willing to forgo in equilibrium r to be equally well-off as in the alternative equilibrium a. Here we follow CCK and define λ^c as the fraction of consumption which has to be given to the representative agent in equilibrium r in order to be as well-off as in equilibrium a, i.e.,

$$V_{at}=\mathbb{E}_t\left\{\sum_{s=0}^{\infty}\beta^s\frac{1}{1-\eta}[(1+\lambda^c)c_{rt+1}]^{\alpha(1-\eta)}(1-n_{rt+s})^{(1-\alpha)(1-\eta)}\right\}=(1+\lambda^c)^{\alpha(1-\eta)}V_{rt}.$$

Solving for λ^c yields

$$\lambda^c=\left[\frac{V_{at}}{V_{rt}}\right]^{\frac{1}{\alpha(1-\eta)}}-1. \tag{8}$$

The superscript c shall remind the reader that this measure is conditioned on the initial point $(k_t,\ln A_t)$.

In order to evaluate λ^c, we will use the second-order approximation of V_{it} at the stationary solution:

$$V_{it}\simeq V_i+H_w^{V_i}(\mathbf{w}_t-\mathbf{w})+\tfrac{1}{2}(\mathbf{w}_t-\mathbf{w})^T H_{ww}^{V_i}(\mathbf{w}_t-\mathbf{w})+\tfrac{1}{2}H_{\sigma\sigma}^{V_i}. \tag{9}$$

To get this solution, we must add an equation for the variable V_{it} to the system (4) and define an initial condition $(\mathbf{w}_t-\mathbf{w})$. The first requirement is easily met, because the infinite sum (7) has a recursive representation:

$$V_{it}=\frac{1}{1-\eta}c_t^{\alpha(1-\eta)}(1-n_t)^{(1-\alpha)(1-\eta)}+\beta\mathbb{E}_tV_{it+1}. \tag{10}$$

The answer to the second question derives from our ultimate goal. We want to compare an economy without uncertainty and, thus, without TFP shocks to one with a given amount of risk as specified by the choice of τ^2. Therefore, we may assume that both economies start at the stationary solution and that the one without shocks stays there for ever. This gives

$$V_r=\frac{1}{(1-\beta)(1-\eta)}c^{\alpha(1-\eta)}(1-n)^{(1-\alpha)(1-\eta)}$$

for the reference solution. The second-order approximation of V_{at} follows from (6b) for $\mathbf{w}_t-\mathbf{w}=\mathbf{0}_{n(w)}$ and is given by

$$V_{at}\simeq V_r+\tfrac{1}{2}H_{\sigma\sigma}^{V_a},$$

where $H_{\sigma\sigma}^{V_a}$ is the element of the vector $H_{\sigma\sigma}^{y}$ which refers to the variable V_{at}. In the model of Section 2 the non-zero mean of ϵ_{t+1} increases the absolute value of $H_{\sigma\sigma}^{V_a}<0$, so that the effect of uncertainty dominates the mean effect even for small values of τ.

Unconditional measure. Schmitt-Grohé and Uribe (2006) also develop a measure of unconditional welfare gains or losses. They are not very specific about the details but inspection of their Matlab programs allows one to recover the general procedure. We start by integrating both sides of equation (7) with respect to the distribution of the state variables $\mathbf{w}_t$. This gives

$$V_{it}^u := \mathbb{E}\left\{\mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s\frac{1}{1-\eta}c_{it+s}^{\alpha(1-\eta)}(1-n_{it+s})^{(1-\alpha)(1-\eta)}\right]\right\}, \tag{11}$$

where $\mathbb{E}$ denotes unconditional expectations. As before, let λ^u (u for unconditional) denote the fraction of consumption c_{rt+s} the representative household would require to be as well-off in equilibrium r as in equilibrium a:

$$\begin{aligned} V_{at}^u &= \mathbb{E}\left\{\mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s\frac{1}{1-\eta}\left((1+\lambda^u)c_{rt+s}\right)^{\alpha(1-\eta)}(1-n_{rt+s})^{(1-\alpha)(1-\eta)}\right]\right\}, \\ &= (1+\lambda^u)^{\alpha(1-\eta)}V_{rt}^u. \end{aligned}$$

Accordingly, λ^u is given by

$$\lambda^u = \left[\frac{V_{at}^u}{V_{rt}^u}\right]^{\frac{1}{\alpha(1-\eta)}} - 1. \tag{12}$$

In order to obtain the unconditional moments, we take unconditional expectations $\mathbb{E}$ on both sides of equation (6b):

$$\mathbb{E}\bar{\mathbf{y}}_t = H_w^y\mathbb{E}\bar{\mathbf{w}}_t + \tfrac{1}{2}\mathbb{E}\left[I_{n(y)}\otimes\bar{\mathbf{w}}_t^T\right]H_{ww}^y\bar{\mathbf{w}}_t + \tfrac{1}{2}H_{\sigma\sigma}^y,$$

where the bar denotes deviations from the stationary solution, i.e., $\bar{\mathbf{y}}_t := \mathbf{y}_t - \mathbf{y}$ and $\bar{\mathbf{w}}_t := \mathbf{w}_t - \mathbf{w}$. Employing the trace operator to the second term on the right-hand side of this equation and letting Γ^w denote the covariance matrix of the states gives

$$\mathbb{E}\bar{\mathbf{y}}_t = H_w^y\mathbb{E}\bar{\mathbf{w}}_t + \tfrac{1}{2}\begin{bmatrix}\mathrm{tr}\left\{H_{ww}^{y_1}\Gamma^w\right\}\\ \vdots \\ \mathrm{tr}\left\{H^{y_{n(y)}}\Gamma^w\right\}\end{bmatrix} + \tfrac{1}{2}H_{\sigma\sigma}^y. \tag{13}$$

The unconditional expectation of the states follows from stacking equations (6a) and (4c),

$$\begin{aligned} &\bar{\mathbf{w}}_{t+1} = \tilde{H}_w^w\bar{\mathbf{w}}_t + \tfrac{1}{2}\left[I_{n(w)}\otimes\bar{\mathbf{w}}_{t+1}^T\right]\tilde{H}_{ww}^w\bar{\mathbf{w}}_t + \tfrac{1}{2}\tilde{H}_{\sigma\sigma}^w + \tilde{\boldsymbol{\nu}}_{t+1}, \\ &\tilde{H}_w^w := \begin{bmatrix} H_w^x & \\ \mathbf{0}_{n(z)\times n(x)} & R\end{bmatrix}, \; \tilde{H}_{ww}^w := \begin{bmatrix} H_{ww}^x \\ \mathbf{0}_{n(z)n(w)\times n(w)}\end{bmatrix}, \; \tilde{\boldsymbol{\nu}}_{t+1} := \begin{bmatrix}\mathbf{0}_{n(x)\times 1} \\ \boldsymbol{\mu} + \Omega\boldsymbol{\nu}_{t+1}\end{bmatrix}, \end{aligned}$$

where $\boldsymbol{\mu} = \boldsymbol{\mu}(\sigma = 1)$ and taking expectations on both sides. The result is a linear system in the unknown vector $\mathbb{E}\bar{\mathbf{w}}_t$

$$\left[I_{n(w)} - \tilde{H}_w^w\right]\mathbb{E}\bar{\mathbf{w}}_t = \mathbb{E}\tilde{\boldsymbol{\nu}}_{t+1} + \tfrac{1}{2}\begin{bmatrix}\mathrm{tr}\left\{\tilde{H}_{ww}^{w_1}\Gamma^w\right\}\\ \vdots \\ \mathrm{tr}\left\{\tilde{H}_{ww}^{w_{n(w)}}\Gamma^w\right\}\end{bmatrix} + \tfrac{1}{2}\tilde{H}_{\sigma\sigma}^w. \tag{14}$$

In the last step we determine the covariance matrix Γ^w from the linearized solution for the vector of states:

$$\bar{\mathbf{w}}_{t+1} = \tilde{H}^w_w \bar{\mathbf{w}}_t + \tilde{\boldsymbol{\nu}}_{t+1}.$$

The matrix Γ^w solves the discrete Lyapunov equation[9]

$$\Gamma^w = \tilde{H}^w_w \Gamma^w \left(\tilde{H}^w_w\right)^T + \Sigma_{\tilde{\nu}}, \ \Sigma_{\tilde{\nu}} = \begin{bmatrix} \mathbf{0}_{n(x)\times n(x)} & \mathbf{0}_{n(x)\times n(z)} \\ \mathbf{0}_{n(z)\times n(x)} & \Omega\Omega^T \end{bmatrix}. \tag{15}$$

Note that the Matlab program of Schmitt-Grohé and Uribe (2004b) does not consider the term $\mathbb{E}\tilde{\boldsymbol{\nu}}_{t+1} = [\mathbf{0}_{1\times n(x)}, \boldsymbol{\mu}^T]^T$, so that the unconditional measure λ^u is biased by i) disregarding this effect and by ii) disregarding the effect of $\boldsymbol{\mu}_{\sigma\sigma}$ on the vector $\tilde{H}^w_{\sigma\sigma}$.

Mean and Fluctuations Effect. CCK decompose the welfare effect into two components. The mean effect ω^m reflects the optimal response of labor supply and capital accumulation to the productivity shock. It may be increasing in the standard deviation of the productivity shock τ. The second component ω^u captures risk-aversion, and, thus, will always decrease in the amount of uncertainty as measured by τ. CCK define the mean effect as the fraction of consumption required by agents living in the stationary environment to be equally well-off as agents who enjoy the expected value of consumption and leisure obtained from living in the stochastic environment. Let $\tilde{c}$ and $\tilde{n}$ denote the expected values of consumption and hours obtained from the solution of equation (13). Then, the mean effect is defined by

$$(1+\omega^m)^{\alpha(1-\eta)} V_r = \tilde{V} := \sum_{s=0}^{\infty} \beta^s \tilde{c}^{\alpha(1-\eta)} (1-\tilde{n})^{(1-\alpha)(1-\eta)}. \tag{16}$$

Accordingly, the fluctuations effect solves the equation

$$V^u_{at} = (1+\omega^f)^{\alpha(1-\eta)} \tilde{V}, \tag{17}$$

so that agents in the stochastic environment enjoy the same unconditional expected lifetime utility as those agents being provided with a steady stream of consumption equal to $(1+\omega^f)\tilde{c}$ and working a constant fraction of $\tilde{n}$ hours. Combining

$$V^u_{at} = (1+\lambda^u) V_t$$

with the previous two equations gives

$$(1+\omega^m)(1+\omega^f) = (1+\lambda^u),$$

so that the welfare effect λ^u is approximately equal to the sum of the mean and the fluctuations effect.

[9] For the derivation of this equation see, e.g., Hamilton (1994), p. 264f or Lütkepohl (2005) p. 26f.

3.3 Mean Weighted Residuals Methods

In order to check the logic underlying our assumptions in (4c)-(4f) and the accuracy of our perturbation solution, we also solve the model with two variants of a mean weighted residuals method (see Appendix C for a detailed description). In particular, we use a finite element method (FEM)[10] and a Chebyshev-Galerkin method (CGM)[11] to approximate the solutions $n(k_t, \ln A_t)$ for hours and $V(k_t, \ln A_t)$ for the value function. In both of these methods it is easy to implement the assumption on the mean of the innovations in equation (2): In the step where conditional expectations with respect to the distribution of ϵ_{t+1} have to be taken, we only have to adjust the mean.

The finite element method employs a grid $\Gamma \subset \mathbb{R}^2$ over the space of the state variables. Between the grid points cubic $\mathscr{C}^2$ splines approximate the respective function. We determine the function values at the grid points such that the residuals from (4a) vanish. In order to evaluate expectations with respect to the $N(\mu, \Omega)$-distributed vector of innovations $\boldsymbol{\epsilon}_{t+1}$ we employ Gauss-Hermite quadrature.

The Chebyshev-Galerkin method employs a tensor product base of Chebyshev polynomials to approximate the solution for hours and the value function. For instance, the solution for hours is approximated by

$$\hat{h}(k_t, z_t) := \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \phi_{i,j} T_{i-1}(\psi_1(k_t)) T_{j-1}(\psi_2(z_t)), \quad z_t := \ln A_t$$

where T_l denotes the l-th order Chebychev polynomial, $d_1, d_2 \in \mathbb{N}$ are the degrees of approximation, and ψ_i, $i = 1, 2$ are bijections between $[\underline{k}, \bar{k}]$ and $[\underline{z}, \bar{z}]$, respectively, into the domain of Chebyshev polynomials $[-1, 1]$. The $d_1 d_2$ coefficients $\phi_{i,j}$ solve a system of non-linear equations which requires that weighted sums of the residuals of (4a) vanish.

4 Results

In this section we present the welfare gains and losses computed for the model from Section 2 and discuss the accuracy of these measures. Our results rest on the benchmark calibration of the model in Table 1. In order to study the effects of different degrees of risk-aversion and uncertainty, we follow CCK and solve the model for ten different values of the parameter η and five different values of the parameter τ. As in CCK, we report the measures in percent of income, i.e., instead of λ we report $\tilde{\lambda} := \lambda(c/y)$, where c and y denote, respectively, the stationary value of consumption and income. Our results are summarized in Table 2 and illustrated in Figure 1. The second column in Table 2 designates the measure, where the superscript c (u) refers to the conditional (unconditional) measure and the subscript 0 indicates that we have neglected the effect of the shifting mean on the solution. Solutions based on a second-order perturbation are labeled “pert” in the third column of Table 2 while those obtained from the weighted residuals methods are labeled “proj”.

[10]See, e.g., McGrattan (1995).
[11]See, e.g., Judd (1992) and for a textbook presentation Heer and Maußner (2009), Chapter 6.

4.1 Welfare Measures

First, consider Figure 1. In its panel (a) it displays the conditional and in its panel (b) the unconditional welfare measures from our extension of the perturbation solution of the model. Panels (c) and (d) show our attempts to reproduce Figure 2 of CCK. Finally, panel (e) displays the conditional measure from our Galerkin-Chebshev solution while panel (f) shows the unconditional measure from our finite elements solution.

First, note that both the conditional and the unconditional measures from our method indicate welfare gains from removing the business cycle for all combinations of risk-aversion and the amount of risk as parameterized by η and τ, respectively. Hence, the fluctuations effect always dominates the mean effect, so that the representative agent living in the stationary and certain environment would not want to live in the stochastic economy, i.e., $\tilde{\lambda}^i < 0$. This is quite different from the results reported in Figure 2 of CCK. In panels (c) and (d) we try to reproduce their Figure 2 by ignoring the effect of $\mu_{\sigma\sigma}$ on the solution of the model. However, we computed the unconditional measure from equation (14) by assuming $\mathbb{E}\tilde{\boldsymbol{\nu}}_{t+1} = -\tau^2/(2(1+\rho))$. The graphs show that both measures are positive for values of η smaller than (about) 4 (Panel (c)) and 5.5 (Panel (d)). In Figure 2 of CCK the graphs cut the abscissa at a value of η slightly larger than 5. In addition, the starting and endpoints of each graph in Panel (d) are close to those in Figure 2 of CCK, so that we suspect that their Figure 2 reports the unconditional welfare measure but does not perturb the mean of the innovations in equation (2).[12]

Thus, as a first result, we note that a proper implementation of the mean preserving spread property requires to perturb both the mean and the variance of the innovations. Otherwise one overstates the mean relative to the fluctuations effect. This can be seen by comparing the lines labeled ω^m and ω_0^m in Table 2. They show, respectively, the mean effect computed from our proposed method and the mean effect, if we disregard the effect of $\mu_{\sigma\sigma}$ on $H_{\sigma\sigma}$. The latter overstates the former by at least 56.7 percent and at most by 111.7 percent.

Table 2: Welfare Measures

η	M	S	τ 0.003	0.007	0.011	0.015	0.019
1.0	$\tilde{\lambda}^c$	pert	−0.000951	−0.005179	−0.012788	−0.023777	−0.038146
	$\tilde{\lambda}^c_0$	pert	0.002888	0.015728	0.038844	0.072246	0.115949
	$\tilde{\lambda}^c$	proj	−0.000951	−0.005178	−0.012784	−0.023760	−0.038067
	$\tilde{\lambda}^u$	pert	−0.001130	−0.006149	−0.015184	−0.028233	−0.045293
	$\tilde{\lambda}^u_0$	pert	0.004233	0.023050	0.056931	0.105899	0.169982
	$\tilde{\lambda}^u$	proj	−0.001091	−0.006059	−0.014953	−0.024757	−0.019034
	$\tilde{\omega}^m$	pert	0.004801	0.026135	0.064527	0.119961	0.192414
	$\tilde{\omega}^m_0$	pert	0.010163	0.055317	0.136540	0.253739	0.406778

continued on next page

[12]We also computed the unconditional measure $\tilde{\lambda}^u$ by assuming $\mathbb{E}\tilde{\boldsymbol{\nu}}_{t+1} = 0$ in equation (14). The point of intersection of the graphs with the abscissa remains about the same, yet the size of the effects becomes smaller. For instance, we find $\tilde{\lambda}^u(\eta = 1, \tau = 0.019) = 0.126$ and $\tilde{\lambda}^u(\eta = 10, \tau = 0.019) = -0.072$ while Figure 2 of CCK clearly shows a value above 0.15 and below -0.1, respectively.

η	M	S	τ				
			0.003	0.007	0.011	0.015	0.019
	$\tilde{\omega}^u$	proj	0.004915	0.026400	0.064924	0.120071	0.190095
2.0	$\tilde{\lambda}^c$	pert	−0.002074	−0.011291	−0.027878	−0.051828	−0.083132
	$\tilde{\lambda}^c_0$	pert	0.001766	0.009613	0.023742	0.044157	0.070865
	$\tilde{\lambda}^c$	proj	−0.002074	−0.011290	−0.027873	−0.051805	−0.083022
	$\tilde{\lambda}^u$	pert	−0.002305	−0.012551	−0.030988	−0.057608	−0.092399
	$\tilde{\lambda}^u_0$	pert	0.003057	0.016646	0.041115	0.076477	0.122755
	$\tilde{\lambda}^u$	proj	−0.002264	−0.012454	−0.030701	−0.053140	−0.063826
	$\tilde{\omega}^m$	pert	0.004870	0.026518	0.065488	0.121788	0.195431
	$\tilde{\omega}^m_0$	pert	0.010233	0.055715	0.137594	0.255888	0.410623
	$\tilde{\omega}^u$	proj	0.004988	0.026788	0.065874	0.121655	0.192116
3.0	$\tilde{\lambda}^c$	pert	−0.003079	−0.016761	−0.041377	−0.076909	−0.123331
	$\tilde{\lambda}^c_0$	pert	0.000761	0.004141	0.010227	0.019020	0.030520
	$\tilde{\lambda}^c$	proj	−0.003079	−0.016760	−0.041374	−0.076890	−0.123206
	$\tilde{\lambda}^u$	pert	−0.003268	−0.017788	−0.043911	−0.081618	−0.130878
	$\tilde{\lambda}^u_0$	pert	0.002095	0.011406	0.028171	0.052399	0.084102
	$\tilde{\lambda}^u$	proj	−0.003223	−0.017684	−0.043550	−0.076115	−0.092269
	$\tilde{\omega}^m$	pert	0.005168	0.028135	0.069479	0.129205	0.207320
	$\tilde{\omega}^m_0$	pert	0.010530	0.057332	0.141581	0.263286	0.422464
	$\tilde{\omega}^u$	proj	0.005288	0.028409	0.069841	0.128765	0.202015
4.0	$\tilde{\lambda}^c$	pert	−0.004024	−0.021903	−0.054063	−0.100466	−0.161058
	$\tilde{\lambda}^c_0$	pert	−0.000185	−0.001005	−0.002481	−0.004613	−0.007401
	$\tilde{\lambda}^c$	proj	−0.004024	−0.021904	−0.054068	−0.100468	−0.160946
	$\tilde{\lambda}^u$	pert	−0.004110	−0.022371	−0.055218	−0.102611	−0.164494
	$\tilde{\lambda}^u_0$	pert	0.001252	0.006818	0.016839	0.031319	0.050262
	$\tilde{\lambda}^u$	proj	−0.004063	−0.022262	−0.054749	−0.095558	−0.119368
	$\tilde{\omega}^m$	pert	0.005585	0.030407	0.075086	0.139622	0.224013
	$\tilde{\omega}^m_0$	pert	0.010947	0.059602	0.147181	0.273682	0.439104
	$\tilde{\omega}^u$	proj	0.005708	0.030683	0.075406	0.138769	0.217019
5.0	$\tilde{\lambda}^c$	pert	−0.004930	−0.026832	−0.066216	−0.123017	−0.197137
	$\tilde{\lambda}^c_0$	pert	−0.001091	−0.005938	−0.014661	−0.027257	−0.043721
	$\tilde{\lambda}^c$	proj	−0.004930	−0.026835	−0.066237	−0.123062	−0.197083
	$\tilde{\lambda}^u$	pert	−0.004867	−0.026488	−0.065367	−0.121441	−0.194615
	$\tilde{\lambda}^u_0$	pert	0.000495	0.002697	0.006660	0.012385	0.019874
	$\tilde{\lambda}^u$	proj	−0.004816	−0.026375	−0.064726	−0.112567	−0.141774
	$\tilde{\omega}^m$	pert	0.006086	0.033132	0.081810	0.152112	0.244026
	$\tilde{\omega}^m_0$	pert	0.011448	0.062326	0.153898	0.286149	0.459055
	$\tilde{\omega}^u$	proj	0.006211	0.033407	0.082052	0.150825	0.235319
6.0	$\tilde{\lambda}^c$	pert	−0.005807	−0.031603	−0.077973	−0.144812	−0.231968
	$\tilde{\lambda}^c_0$	pert	−0.001968	−0.010714	−0.026451	−0.049164	−0.078838
	$\tilde{\lambda}^c$	proj	−0.005808	−0.031611	−0.078019	−0.144930	−0.232027
	$\tilde{\lambda}^u$	pert	−0.005555	−0.030231	−0.074591	−0.138537	−0.221932
	$\tilde{\lambda}^u_0$	pert	−0.000193	−0.001052	−0.002598	−0.004831	−0.007750
	$\tilde{\lambda}^u$	proj	−0.005502	−0.030115	−0.073747	−0.126707	−0.159711
	$\tilde{\omega}^m$	pert	0.006653	0.036220	0.089428	0.166259	0.266684
	$\tilde{\omega}^m_0$	pert	0.012016	0.065413	0.161509	0.300270	0.481648
	$\tilde{\omega}^u$	proj	0.006782	0.036491	0.089585	0.164102	0.255843

continued on next page

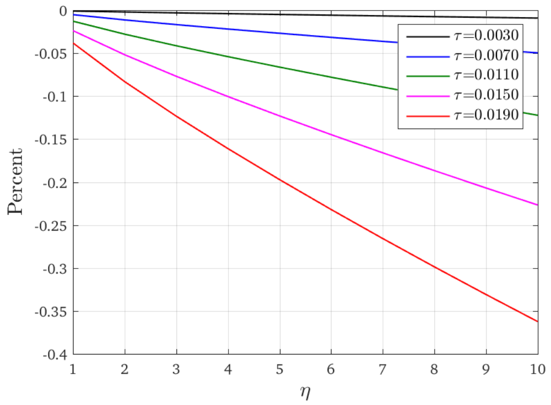
η	M	S	τ				
			0.003	0.007	0.011	0.015	0.019
7.0	$\tilde{\lambda}^c$	pert	−0.006662	−0.036249	−0.089414	−0.166000	−0.265783
	$\tilde{\lambda}^c_0$	pert	−0.002823	−0.015367	−0.037929	−0.070480	−0.112982
	$\tilde{\lambda}^c$	proj	−0.006663	−0.036263	−0.089496	−0.166222	−0.266018
	$\tilde{\lambda}^u$	pert	−0.006186	−0.033658	−0.083029	−0.154163	−0.246867
	$\tilde{\lambda}^u_0$	pert	−0.000824	−0.004485	−0.011075	−0.020589	−0.033026
	$\tilde{\lambda}^u$	proj	−0.006130	−0.033540	−0.081926	−0.139761	−0.174899
	$\tilde{\omega}^m$	pert	0.007278	0.039622	0.097821	0.181840	0.291629
	$\tilde{\omega}^m_0$	pert	0.012641	0.068814	0.169894	0.315823	0.506520
	$\tilde{\omega}^u$	proj	0.007409	0.039887	0.097877	0.179115	0.279194
8.0	$\tilde{\lambda}^c$	pert	−0.007498	−0.040792	−0.100591	−0.186675	−0.298729
	$\tilde{\lambda}^c_0$	pert	−0.003659	−0.019916	−0.049148	−0.091301	−0.146300
	$\tilde{\lambda}^c$	proj	−0.007499	−0.040814	−0.100721	−0.187037	−0.299214
	$\tilde{\lambda}^u$	pert	−0.006765	−0.036807	−0.090777	−0.168496	−0.269706
	$\tilde{\lambda}^u_0$	pert	−0.001404	−0.007642	−0.018865	−0.035066	−0.056234
	$\tilde{\lambda}^u$	proj	−0.006707	−0.036689	−0.089369	−0.151139	−0.186802
	$\tilde{\omega}^m$	pert	0.007956	0.043311	0.106916	0.198718	0.318639
	$\tilde{\omega}^m_0$	pert	0.013319	0.072501	0.178980	0.332671	0.533452
	$\tilde{\omega}^u$	proj	0.008089	0.043564	0.106846	0.195450	0.304857
9.0	$\tilde{\lambda}^c$	pert	−0.008318	−0.045246	−0.111541	−0.206905	−0.330912
	$\tilde{\lambda}^c_0$	pert	−0.004480	−0.024378	−0.060145	−0.111692	−0.178895
	$\tilde{\lambda}^c$	proj	−0.008319	−0.045279	−0.111733	−0.207446	−0.331726
	$\tilde{\lambda}^u$	pert	−0.007299	−0.039707	−0.097905	−0.181665	−0.290658
	$\tilde{\lambda}^u_0$	pert	−0.001938	−0.010548	−0.026038	−0.048390	−0.077581
	$\tilde{\lambda}^u$	proj	−0.007239	−0.039588	−0.096099	−0.160934	−0.195470
	$\tilde{\omega}^m$	pert	0.008683	0.047265	0.116665	0.216804	0.347567
	$\tilde{\omega}^m_0$	pert	0.014046	0.076454	0.188720	0.350725	0.562298
	$\tilde{\omega}^u$	proj	0.008819	0.047504	0.116462	0.213040	0.332751
10.0	$\tilde{\lambda}^c$	pert	−0.009125	−0.049624	−0.122291	−0.226737	−0.362406
	$\tilde{\lambda}^c_0$	pert	−0.005287	−0.028764	−0.070947	−0.131702	−0.210841
	$\tilde{\lambda}^c$	proj	−0.009126	−0.049669	−0.122561	−0.227503	−0.363636
	$\tilde{\lambda}^u$	pert	−0.007791	−0.042378	−0.104466	−0.193771	−0.309884
	$\tilde{\lambda}^u_0$	pert	−0.002430	−0.013228	−0.032646	−0.060658	−0.097224
	$\tilde{\lambda}^u$	proj	−0.007728	−0.042260	−0.102175	−0.169828	−0.200871
	$\tilde{\omega}^m$	pert	0.009457	0.051472	0.127035	0.236033	0.378307
	$\tilde{\omega}^m_0$	pert	0.014819	0.080660	0.199079	0.369920	0.592951
	$\tilde{\omega}^u$	proj	0.009594	0.051690	0.126688	0.231625	0.362831

Notes: $\tilde{\lambda}^c$ and $\tilde{\lambda}^u$ denote, respectively, the conditional and unconditional welfare measure. $\tilde{\omega}^m$ is the mean effect. All measures are in percent of income. The index 0 refers to measures where we have ignored the effect of $\mu_{\sigma\sigma}$ on $H_{\sigma\sigma}$. Lines labeled pert in column S denote the welfare measure from the perturbation solution, while those labeled pert derive from the projection solution.

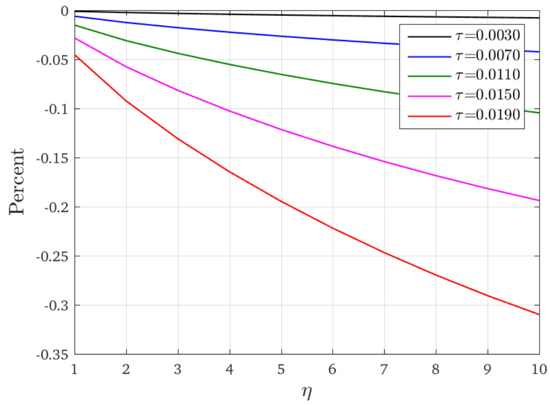
Let us turn now to the size of the effects. As is to be expected, the welfare loss increases with increasing risk-aversion, measured by the parameter η, and with increasing uncertainty, measured by the parameter τ. The welfare gains from removing fluctuation are small. According to Table 2, they range from 0.00051 percent to 0.36 percent of income for the conditional measure and from 0.0011 percent to 0.31 percent for the unconditional

Figure 1: Welfare Measures

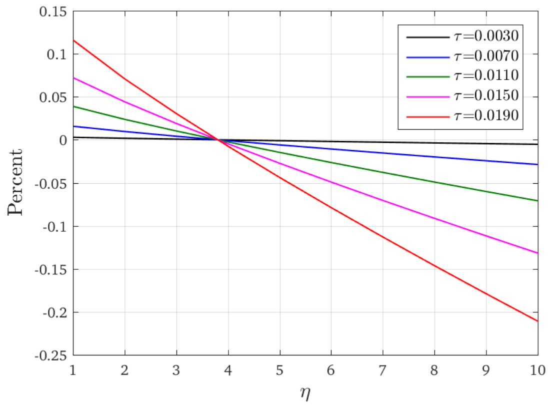
(a) Conditional Measure (Perturbation)



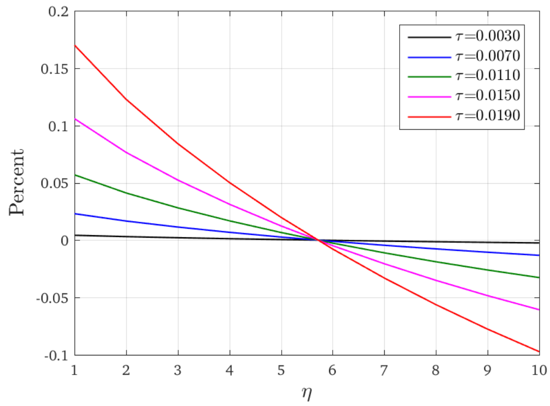
(b) Unconditional Measure (Perturbation)



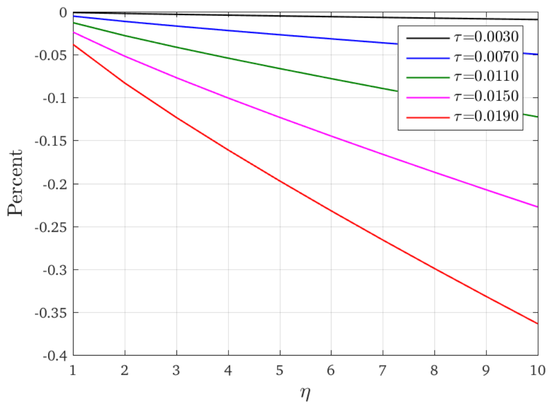
(c) Conditional Measure (CCK) $\mu_{\sigma\sigma} = 0$



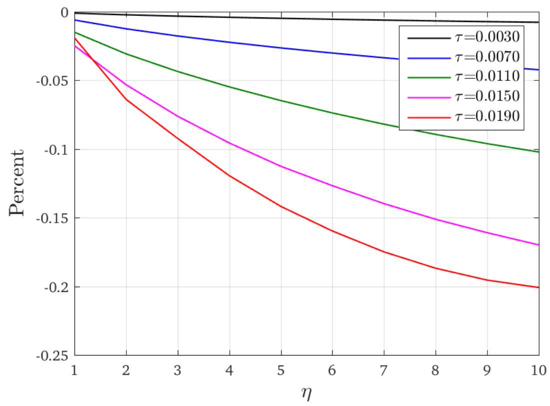
(d) Unconditional Measure (CCK) $\mu_{\sigma\sigma} = 0$



(e) Conditional Measure (CGM)



(f) Unconditional Measure (FEM)



measure. For the latter and in terms of U.S. real per capita income in 2017 of about 52 thousand Dollars (see https://fred.stlouisfed.org) the representative household would be willing to give up between 58 cents and 162.5 Dollars of annual consumption in order to stay within the certain environment. For the common choice of $\tau = 0.007$ and an intermediate value of $\eta = 5$ the figure would be 14 Dollars of annual consumption. This is not far from Lucas's $ 8.5.

Thus, as a second result, we confirm CCK and find small welfare gains from removing the cycle in the benchmark real business cycle model for plausible values of the risk-aversion parameter and the standard deviation of the TFP shock. However, the discrepancy between our proposed solution and the solution which ignores the perturbed mean is large. The percentage deviation $\Delta^i := |(\tilde{\lambda}_0^i - \tilde{\lambda}^i)/\tilde{\lambda}^i|$ lies between 41.8 percent and almost 404 percent for the conditional welfare measure ($i = c$) and between 68.6 percent and 475 percent for the unconditional measure ($i = u$).

The lines labeled $\tilde{\omega}^m$ in Table 2 report the size of the mean effect as computed from equation (16). The size of this effect is increasing in both the risk aversion parameter η and the standard deviation τ and ranges between 0.05 percent of income and 0.38 percent. Relative to the absolute value of the overall effect, the mean effect is large for $\eta = 1$ (log preferences) and about 4.25 times the size of the overall effect, irrespective of the value of τ. For $\eta = 2$ it is still about twice the size of the overall effect. For the remaining values of η between 3 and 10 the mean effect is between 1.6 and 1.17 of the size of the overall effect.

4.2 Accuracy

In order to add proof for the accuracy of our method to implement the mean preserving spread property we also show solutions from our mean weighted residuals methods. We gauge the accuracy of all three solutions of the model with Euler equation residuals. For both functional equations (A.40) and (A.41) defined in Appendix C we compute the percentage increase in consumption that would be required to equate the left-hand side to the right-hand side, if we insert the approximate instead of the true policy function.[13] We compute the residuals on an equal spaced rectangular grid of 100×100 points over the square $[0.85k, 1.15k] \times [-3.7\tau/\sqrt{1-\rho^2}, 3.7\tau/\sqrt{1-\rho^2}]$. The maximum absolute value from this grid is our accuracy measure.

In terms of this measure, and as is well-known from, e.g., Aruoba et al. (2006) and Heer and Maußner (2008), the perturbation solution performs worst. Its Euler residuals for the functional equation (A.41) (which determines the value function required in equation (8)) range from 0.06 percent for $\eta = 1$ and $\tau = 0.003$ to 3.99 percent for $\eta = 10$ and $\tau = 0.019$. For the finite element method they range from 0.00077 percent for $\eta = 1$ and $\tau = 0.003$ to 0.01233 percent for $\eta = 1$ and $\tau = 0.019$. Even more accurate is the Chebyshev-Galerkin method with a minimum Euler equation residual of 2.2×10^{-8} percent for $\eta = 1$ and $\tau = 0.003$ and a maximum value of 7.96×10^{-6} for $\eta = 10$ and $\tau = 0.019$. For this reason

[13] See Christiano and Fisher (2000) for this kind of Euler residuals. The original concept is introduced in Judd and Guu (1997).

we also present the conditional welfare measure computed from this solution in the lines labeled proj in column S of Table 2 and displayed in panel (e) of Figure 1.

The absolute difference between both measures lies between 1.1×10^{-6} percentage points for $\eta = 9$ and $\tau = 0.003$ and 1.2×10^{-3} percentage points for $\eta = 10$ and $\tau = 0.019$. For the latter this difference is equal to about 62 cents of annual U.S. income in 2017. Hence, even though the perturbation solution is quite less accurate outside a small neighborhood of the stationary solution, this imprecision is of almost negligible consequence for the conditional welfare measure.

In order to compare unconditional measures from the perturbation approach to those obtained from the weighted residuals methods we have to resort to simulation, since the distribution of the state variables $\mathbf{w}_t := (k_t, \ln A_t)^T$ is unknown. Our estimate of the unconditional mean of V_t^u from equation (11) is

$$V_{at} = \frac{1}{T} \sum_{s=0}^{T} \hat{V}(k_{t+s}, \ln A_{t+s}),$$

where $\hat{V}(k_t, \ln A_t)$ denotes the approximate FEM solution of the value function. Since the welfare effects for small values of η and τ are of the order of magnitude between 10^{-4} and 10^{-5}, we must choose T very large. To see this, note that the variance of the sample mean $\bar{z}$ of log TFP is equal to

$$\text{var}(\bar{z}) = \frac{1}{T^2} \frac{\tau^2}{1-\rho^2} \left[T + \frac{2}{(1-\rho)^2} \left((T-1)\rho - T\rho^2 + \rho^{T+1} \right) \right].$$

For $T = 5 \times 10^8$ this is still as large as 1.7×10^{-5}. Table 2 displays the means from 15 simulations of size $T = 5 \times 10^8$ each. Compared to the results from the perturbation solution the FEM solution yields smaller effects being close to those from the perturbation solution for $\tau \leq 0.011$. For $\tau = 0.015$ the FEM measures are between 7 and 13 percent smaller than those from the perturbation solution. For the even larger value of $\tau = 0.019$, the FEM measure is between 27 and 58 percent smaller. We trace this difference to the fluctuations effect, since the FEM measure of the unconditional mean effect is close to the mean effect from the perturbation solution. At the maximum, the former exceeds the latter by 2.4 percent for $\eta = 2$ and $\tau = 0.030$) and is 4.3 percent below the latter for ($\eta = 8$ and $\tau = 0.019$). We, thus, conclude that the perturbation approach overestimates the unconditional fluctuations effect for large TFP shocks.

As a third result, we therefore note that our perturbation method delivers reliable estimates of the conditional welfare measure and also for the unconditional measure if the amount of uncertainty in the model as measured by τ is of moderate size.

5 Conclusion

Our paper is inspired by the contribution of Cho, Cooley, and Kim (2015) (CCK) who argue convincingly that the welfare effects of economic fluctuations in production economies consist of two parts. The fluctuations effect captures the fact that risk-averse economic agents would prefer living in a certain environment. The mean effect stems from the optimal adjustment of input factors. Economic agents may benefit from this effect, if the reduced form aggregate production function is convex in the shocks that drive the model. Measuring both effects without distortion requires that the shocks are modeled as mean-preserving spreads.

Introducing mean preserving spreads into a model requires to consider their means as functions of their variance. As a consequence, approximations based on perturbing the variance should adequately and simultaneously also perturb the mean. Otherwise, by comparing different levels of uncertainty, the researcher exogenously introduces shifts of the means and mixes level effects with effects from removing uncertainty. Yet, available perturbation methods and code assumes that the mean of innovations in the vector-autoregressive process for shocks is not perturbed. As a consequence, the impact of the mean preserving spreads is missing in the second-order approximation of the value function, which in turn is required to compute the conditional and unconditional welfare measure. Furthermore, the non-zero means also bias the computation of unconditional moments required to compute the unconditional welfare measure.

The main contribution of this paper, therefore, is to extend the canonical stochastic dynamic equilibrium model to allow for mean preserving spreads. We derive formulas for the computation of second-order approximations of the policy functions and provide the respective Matlab code. We find that our second order solution differs from the standard case and interpret this term as the effect of unwanted shifts in means. We then recompute the welfare effects of removing fluctuations in the benchmark real business cycle model that underlies Figure 2 in CCK. Different form their results, the mean effect never dominates the fluctuations effect. Finally, we solve the model with the aid of two weighted residuals methods and confirm both our method and our revised results. Hence, we conclude that the positive welfare effects of the business cycle reported by CCK for small degrees of risk aversion cannot be traced to a large mean effect, driven by economic agents using economic fluctuations to their advantage. Rather, they result from exogenously introduced shifts in the production function that were not adequately removed.

We, therefore, hope that our paper will clarify and resolve a methodological issue involved in the proper computation of the welfare effects of economic fluctuations.

REFERENCES

Aruoba, S. B., J. Fernández-Villaverde, and J. F. Rubio-Ramírez (2006). Comparing solution methods for dynamic equilibrium economies. *Journal of Economic Dynamics and Control 30*, 2477–2508.

Backus, D. K., P. J. Kehoe, and F. E. Kydland (1992). International real business cycles. *Journal of Political Economy 100*.

Barillas, F., L. P. Hansen, and T. J. Sargent (2009). Doubts or variability? *Journal of Economic Theory 144*(6), 2388 – 2418.

Barlevy, G. (2004, September). The cost of business cycles under endogenous growth. *American Economic Review 04*(4), 964–990.

Barlevy, G. (2005). The costs of business cycles and the benefits of stabilization. *Economic Perspectives 33*(1), 32–49.

Barro, R. J. (2009). Rare disasters, asset prices, and welfare costs. *American Economic Review 99*(1), 243–264.

Cho, J.-O., T. F. Cooley, and H. S. K. Kim (2015). Business cycle uncertainty and economic welfare. *Review of Economic Dynamics 18*, 185–200.

Cho, J.-O., T. F. Cooley, and L. Phaneuf (1997). The welfare cost of nominal wage contracting. *The Review of Economic Studies 64*(3), 465–484.

Christiano, L. J. and J. D. Fisher (2000). Algorithms for solving dynamic models with occasionally binding constraints. *Journal of Economic Dynamics and Control 24*, 1179–1232.

De Santis, M. (2007). Individual consumption risk and the welfare costs of business cycles. *American Economic Review 97*(4), 1488–1506.

Dolmas, J. (1998). Risk preferences and the welfare cost of business cycles. *Review of Economic Dynamics 1*(3), 646 – 676.

Galí, J., M. Gertler, and J. D. López-Salido (2007). Markups, gaps, and the welfare costs of business fluctuations. *Review of Economics and Statistics 89*(1), 44–59.

Gomme, P. and P. Klein (2011). Second-order approximation of dynamic models without the use of tensors. *Journal of Economic Dynamics and Control 35*, 604–615.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hansen, G. D. (1985). Indivisible labor and the business cycle. *Journal of Monetary Economics 16*, 309–327.

Heer, B. and A. Maußner (2008). Computation of business cycle models: A comparison of numerical methods. *Macroeconomic Dynamics 12*, 641–663.

Heer, B. and A. Maußner (2009). *Dynamic General Equilibrium Modelling. 2nd Edition*. Berlin: Springer.

Heer, B. and A. Maußner (2015). The cash-in-advance constraint in monetary growth models with labor market search. *Macroeconomic Dynamics 19*, 144–166.

İmrohoroğlu, A. (1989). Cost of business cycles with indivisibilities and liquidity constraints. *Journal of Political Economy 97*(6), 1364–1383.

Judd, K. L. (1992). Projection methods for solving aggregate growth models. *Journal of Economic Theory 58*, 410–452.

Judd, K. L. and S.-M. Guu (1997). Asymptotic methods for aggregate growth models. *Journal of Economic Dynamics and Control 21*.

Kågström, B. and P. Poromaa (1994). Lapack-style algorithms and software for solving the generalized sylvester equation and estimating the separation between regular matrix pairs. Working Notes UT, CS-94-237, LAPACK.

Krebs, T. (2003). Growth and welfare effects of business cycles in economies with idiosyncratic human capital risk. *Review of Economic Dynamics 6*(4), 846 – 868.

Krusell, P., T. Mukoyama, A. Şahin, and A. A. Smith (2009). Revisiting the welfare effects of eliminating business cycles. *Review of Economic Dynamics 12*(3), 393 – 404.

Lucas, Jr, R. E. (1987). *Models of Business Cycles*. Oxford: Basil Blackwell.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer.

Magnus, J. R. and H. Neudecker (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics* (3rd ed.). Chichester: John Wiley &Sons.

McGrattan, E. R. (1995). Solving the stochastic growth model with a finite element method. *Journal of Economic Dynamics and Control 20*, 19–42.

Obstfeld, M. (1994). Evaluating risky consumption paths: The role of intertemporal substitutability. *European Economic Review 38*(7), 1471 – 1486.

Schmitt-Grohé, S. and M. Uribe (2004a). Optimal simple and implementable monetary and fiscal rules. Working Paper W10253, National Bureau of Economic Research (NBER).

Schmitt-Grohé, S. and M. Uribe (2004b). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control 28*, 755–775.

Schmitt-Grohé, S. and M. Uribe (2006). Optimal simple and implementable monetary and fiscal policy rules: Expanded version. Working Paper W12402, National Bureau of Economic Research (NBER).

Schmitt-Grohé, S. and M. Uribe (2007). Optimal simple and implementable monetary and fiscal rules. *Journal of Monetary Economics 54*, 1702–1725.

Sydsæter, K., A. Strøm, and P. Berck (1999). *Economists' Mathematical Manual, Third Edition*. Berlin: Springer.

Tallarini, T. D. (2000). Risk-sensitive real business cycles. *Journal of Monetary Economics 45*(3), 507 – 532.

Appendix

A Equilibrium conditions of the model

In this part of the Appendix we map the model of Section 2 into the canonical DSGE model of equations (4). The model has one endogenous state variable, the capital stock $\mathbf{x}_t := k_t$, and one shock, the log of TFP $\mathbf{z}_t := \ln A_t$. The not predetermined variables are $\mathbf{y}_t := [y_t, c_t, i_t, n_t, \lambda_t, V_t]^T$. The variable λ_t is equal to the Lagrange multiplier of the constraint

$$0 \le y_t - c_t - (k_{t+1} - (1-\delta)k_t)$$

and introduced for notational convenience. The model's equations $\mathbf{g}$ are given by

$$g^1(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t) = y_t - A_t n_t^{1-\theta} k_t^{\theta}, \tag{A.1a}$$

$$g^2(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t) = \begin{cases} \lambda_t - \alpha c_t^{\alpha(1-\eta)-1}(1-n_t)^{(1-\alpha)(1-\eta)} & \text{for } \eta \neq 1, \\ \lambda_t - \frac{\alpha}{c_t} & \text{for } \eta = 1, \end{cases} \tag{A.1b}$$

$$g^3(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t) = (1-\theta)\frac{y_t}{n_t} - \frac{1-\alpha}{\alpha}\frac{c_t}{1-n_t}, \tag{A.1c}$$

$$g^4(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t) = y_t - c_t - i_t, \tag{A.1d}$$

$$g^5(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t) = k_{t+1} - (1-\delta)k_t - i_t, \tag{A.1e}$$

$$g^6(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t) = 1 - \beta\frac{\lambda_{t+1}}{\lambda_t}\left(1 - \delta + \theta\frac{y_{t+1}}{k_{t+1}}\right), \tag{A.1f}$$

$$g^7(\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t) = \begin{cases} V_t - \frac{1}{1-\eta}c_t^{\alpha(1-\eta)}(1-n_t)^{(1-\alpha)(1-\eta)} - \beta V_{t+1} & \text{for } \eta \neq 1, \\ V_t - \alpha \ln c_t - (1-\alpha)\ln(1-n_t) - \beta V_{t+1} & \text{for } \eta = 1, \end{cases} \tag{A.1g}$$

Equation (A.1a) is the production function, equation (A.1b) is the first-order condition with respect to consumption c_t, equation (A.1c) is the first-order condition for labor supply, equation (A.1d) is the resource constraint, equation (A.1e) is the law of motion of the capital stock, equation (A.1f) is the Euler equation for capital accumulation, and (A.1g) is the recursive definition of expected life-time utility. Introducing the parameter σ into the model in order to perturb the volatility τ of shocks, the mean of the innovations in the AR(1)-process of the log of the TFP shock is given by

$$\mu(\sigma) := -\frac{(\sigma\tau)^2}{2(1+\rho)} \tag{A.2}$$

so that – as required by (4d) - (4f) –

$$\mu(\sigma = 0) = 0,$$

$$\mu_\sigma(\sigma = 0) = -\sigma\frac{\tau^2}{(1+\rho)}\bigg|_{\sigma=0} = 0,$$

$$\mu_{\sigma\sigma}(\sigma = 0) = -\frac{\tau^2}{(1+\rho)} < 0.$$

The stationary solution of the model at $\sigma = 0$ and $\mathbb{E}Z = 1$ solves the system of equations:

$$\frac{y}{k} = \frac{1-\beta(1-\delta)}{\beta\theta}, \tag{A.3a}$$

$$\frac{c}{k} = \frac{y}{k} - \delta, \tag{A.3b}$$

$$\frac{n}{1-n} = \alpha\frac{1-\theta}{1-\alpha}\frac{y/k}{c/k}. \tag{A.3c}$$

Given the solution for n, the levels of output, consumption, and the capital follow from the solutions of y/k and c/k. Investment follows from the stationary version of equation (A.1e) and is equal to $i = \delta k$. The stationary version of equation (A.1g) implies

$$V = \begin{cases} \frac{c^{\alpha(1-\eta)}(1-n)^{(1-\alpha)(1-\eta)}}{(1-\beta)(1-\eta)} & \text{for } \eta \neq 1, \\ \frac{\alpha \ln c + (1-\alpha)\ln(1-n)}{1-\beta} & \text{for } \eta = 1. \end{cases}$$

B Computation of second-order coefficients for models with non-zero means

In this section we extend the computation of second-order coefficient matrices H^x_{ww}, H^y_{ww}, $H^x_{\sigma\sigma}$, and $H^y_{\sigma\sigma}$ in the solutions (6) to the case of also perturbing the means of the innovations $\boldsymbol{\epsilon}_{t+1}$ as defined in (4d)-(4f). We follow Gomme and Klein (2011) and employ the chain rule for Hessian matrices of Magnus and Neudecker (1999). Though the effect of assumption about the mean will show up only in the vectors $H^w_{\sigma\sigma}$ and $H^y_{\sigma\sigma}$, we also derive the linear system whose solution yields H^w_{ww} and H^y_{ww}. The reason is that we distinguish between endogenous states $\mathbf{x}_t$ and shocks $\mathbf{z}_t$ while Gomme and Klein (2011) merge both in the vector $\mathbf{w}_t$. The next subsection introduces some notation and presents the chain rules. Section B.2 derives the linear systems of equations whose solutions are the matrices of second-order coefficients.

B.1 Second-order derivatives of composite functions

Let $\mathbf{f} = (f^1, \dots, f^m) : \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{g} = (\mathbf{g}^1, \dots, \mathbf{g}^l) : \mathbb{R}^m \to \mathbb{R}^l$ denote vector valued functions and let $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$, and $\mathbf{z} \in \mathbb{R}^l$ denote real vectors. The composite function $\mathbf{h}(\mathbf{x}) := \mathbf{g} \circ \mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^l$ is defined by

$$\mathbf{h}(\mathbf{x}) := (\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \begin{bmatrix} g^1(f^1(\mathbf{x}), f^2(\mathbf{x}), \dots, f^m(\mathbf{x})) \\ g^2(f^1(\mathbf{x}), f^2(\mathbf{x}), \dots, f^m(\mathbf{x})) \\ \vdots \\ g^l(f^1(\mathbf{x}), f^2(\mathbf{x}), \dots, f^m(\mathbf{x})) \end{bmatrix}. \tag{A.4}$$

The Jacobian of this function with respect to $\mathbf{x}$, denoted by $\mathbf{h}_x = (h_{i,j})_{i=1,\dots,l;j=1,\dots,n} \in \mathbb{R}^{l\times n}$ has the typical element

$$h_{ij} := \sum_{s=1}^{m} \frac{\partial g^i}{\partial y_s}(\mathbf{f}(\mathbf{x}))\frac{\partial f^s}{\partial x_j}(\mathbf{x}),\; i = 1, 2, \dots l; j = 1, 2, \dots, n. \tag{A.5}$$

The typical element of the Hessian matrix $\mathbf{h}_{xx} = (\tilde{h}_{qk})_{q=1,\dots,nl,k=1,\dots,m}$ is obtained by differentiating (A.5) with respect to x_k and is given by

$$\tilde{h}_{(i-1)n+j,k} := \sum_{s=1}^{m} \frac{\partial g^i}{\partial y_s}(\mathbf{f}(\mathbf{x}))\frac{\partial^2 f^s}{x_j x_k}(\mathbf{x}) + \sum_{s=1}^{m}\sum_{r=1}^{m} \frac{\partial f^s}{\partial x_j}(\mathbf{x})\frac{\partial^2 g^i}{\partial y_s \partial y_r}(\mathbf{f}(\mathbf{x}))\frac{\partial f^r}{\partial x_k}(\mathbf{x}),$$
$$i = 1,2,\dots,l; j,k = 1,2,\dots,n. \tag{A.6}$$

Thus, the matrix $\mathbf{h}_{xx}$ is of dimension $nl \times n$, where block $i = 1,\dots,l$ of size $n \times n$ represents the Hessian matrix of the map g^i with respect to the vector $\mathbf{x}$.

The chain rule of Magnus and Neudecker (1999), Theorem 9, shows that $\mathbf{h}_{xx}$ as in (A.6) can be written in terms of the Jacobian and Hessian matrices of $\mathbf{g}$ and $\mathbf{f}$ by:

$$\mathbf{h}_{xx} = \left(I_l \otimes \mathbf{f}_x^T\right)\mathbf{g}_{yy}\mathbf{f}_x + \left(\mathbf{g}_y \otimes I_n\right)\mathbf{f}_{xx}. \tag{A.7}$$

Here $\mathbf{g}_y = (g_{ij})_{i=1,\dots,l;j=1,\dots,m}$ denotes the Jacobian matrix of $\mathbf{g}$ with typical element $g_{ij} := \partial g^i/\partial y_j(\mathbf{f}(\mathbf{x}))$. Its Hessian $\mathbf{g}_{yy} = (\tilde{g}_{qk})_{q=1,\dots,lm;j=1,\dots,m}$ consists of l blocks of size $m \times m$ with typical element $g_{(i-1)l+k,j} := \partial^2 g^i/\partial y_j \partial y_k(\mathbf{f}(\mathbf{x}))$. The Jacobian and Hessian of $\mathbf{f}$ with respect to the vector $\mathbf{x}$, denoted by $\mathbf{f}_x$ and $\mathbf{f}_{xx}$, respectively, are defined analogously.

B.2 Computation of second-order coefficients

Preliminaries. In a first step we put the canonical model (4) into the framework of Section B.1. In the vector

$$\mathbf{s}_t = \left[\mathbf{x}_{t+1}, \mathbf{z}_{t+1}, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t\right] \in \mathbb{R}^{2(n(x)+n(z)+n(y))}$$

we gather the arguments of the vector valued function $\mathbf{g}$ so that (4a) can be written as

$$\mathbf{0}_{(n(x)+n(y))\times 1} = \mathbb{E}_t \mathbf{g}(\mathbf{s}_t), \tag{A.8a}$$

We can now employ (5) and (4c) to define the maps

$$\begin{aligned}\mathbf{f} :& \mathbb{R}^{n(w)+1} \to \mathbb{R}^{n(s)} \quad, (\mathbf{w}_t, \sigma) \mapsto \mathbf{f}(\mathbf{w}_t, \sigma),\\ \mathbf{u} :& \mathbb{R}^{n(w)+1} \to \mathbb{R}^{n(w)+1}, (\mathbf{w}_t, \sigma) \mapsto \mathbf{u}(\mathbf{w}_t, \sigma)\end{aligned}$$

by

$$\mathbf{f}(\mathbf{w}_t, \sigma) := \begin{bmatrix} \mathbf{h}^x(\mathbf{w}_t, \sigma) \\ \Pi \mathbf{z}_t + \boldsymbol{\mu}(\sigma) + \sigma \Omega \boldsymbol{\nu}_{t+1} \\ \mathbf{h}^y\left(\mathbf{u}(\mathbf{w}_t, \sigma)^T\right) \\ \mathbf{x}_t \\ \mathbf{z}_t \\ \mathbf{h}^y(\mathbf{w}_t, \sigma) \end{bmatrix}, \tag{A.8b}$$

$$\mathbf{u}(\mathbf{w}_t, \sigma) := \begin{bmatrix} \mathbf{h}^x(\mathbf{w}_t, \sigma) \\ \Pi \mathbf{z}_t + \boldsymbol{\mu}(\sigma) + \sigma \Omega \boldsymbol{\nu}_{t+1} \\ \sigma \end{bmatrix}. \tag{A.8c}$$

Therefore, the first set of equilibrium conditions (4a) is the composite function $\mathbb{E}_t(\mathbf{g} \circ \mathbf{f})(\mathbf{w}_t, \sigma)$.

In order to simplify the notation in the upcoming derivations we additionally introduce

$$\underline{\mathbf{u}}(\mathbf{w}_t, \sigma) := \begin{bmatrix} \mathbf{h}^x(\mathbf{w}_t, \sigma) \\ \Pi \mathbf{z}_t + \sigma \Omega \boldsymbol{\epsilon}_{t+1} \end{bmatrix} \tag{A.9}$$

for the first $n(w)$ component functions of $\mathbf{u}$ so that

$$\mathbf{u}(\mathbf{w}_t, \sigma) = \begin{bmatrix} \underline{\mathbf{u}}(\mathbf{w}_t, \sigma) \\ \sigma \end{bmatrix}.$$

We already note that the Jacobian $\mathbf{u}_w$ of $\mathbf{u}$ with respect only to $\mathbf{w}_t$ satisfies

$$\mathbf{u}_w = \begin{bmatrix} \underline{\mathbf{u}}_w \\ \mathbf{0}_{1 \times n(w)} \end{bmatrix}. \tag{A.10}$$

Matrices for the Vector of States. Differentiating $\mathbb{E}_t(\mathbf{g} \circ \mathbf{f})$ twice with respect to $\mathbf{w}_t$ and evaluating the result at the point $\mathbf{s}$ yields a system of equations in the coefficients of the matrices $\mathbf{h}^x_{ww}$ and $\mathbf{h}^y_{ww}$. To find this system, we employ (A.7). The Hessian matrix of $\mathbf{g} \circ \mathbf{f}$ with respect to $\mathbf{w}_t$, $(\mathbf{g} \circ \mathbf{f})_{ww}$, is a matrix of size $(n(x) + n(y))n(w) \times n(w)$. Thus:

$$\mathbf{0}_{(n(x)+n(y))n(w) \times n(w)} = (\mathbf{g} \circ \mathbf{f})_{ww} = \left(I_{n(x)+n(y)} \otimes \mathbf{f}_w^T\right) \mathbf{g}_{ss} \mathbf{f}_w + \left(\mathbf{g}_s \otimes I_{n(w)}\right) \mathbf{f}_{ww}. \tag{A.11}$$

In this expression, the matrices $\mathbf{g}_s$ and $\mathbf{g}_{ss}$ are, respectively, the Jacobian matrix and the Hessian matrix of the system of equations with respect to the $2(n(x)+n(z)+n(y))$ variables of the model. The Jacobian $\mathbf{f}_w$ and the Hessian $\mathbf{f}_{ww}$ follow from (A.8b) and (A.8c) via differentiation with respect to $\mathbf{w}_t$. For the Jacobian matrix we get (remember the partition $\mathbf{w} = \left[\mathbf{x}^T, \mathbf{z}^T\right]^T$):

$$\mathbf{f}_w = \begin{bmatrix} \mathbf{h}^x_x & \mathbf{h}^x_z \\ \mathbf{0}_{n(z) \times n(x)} & \Pi \\ \mathbf{h}^y_x \mathbf{h}^x_x & \mathbf{h}^y_x \mathbf{h}^x_z + \mathbf{h}^y_z \Pi \\ I_{n(x)} & \mathbf{0}_{n(x) \times n(z)} \\ \mathbf{0}_{n(z) \times n(x)} & I_{n(z)} \\ \mathbf{h}^y_x & \mathbf{h}^y_z \end{bmatrix}, \tag{A.12}$$

where the third line follows from

$$(\mathbf{h}^y \circ \mathbf{u})_w = \mathbf{h}^y_u \mathbf{u}_w = \begin{bmatrix} \mathbf{h}^y_x & \mathbf{h}^y_z & \mathbf{h}^y_\sigma \end{bmatrix} \begin{bmatrix} \mathbf{h}^x_x & \mathbf{h}^x_z \\ \mathbf{0}_{n(z) \times n(x)} & \Pi \\ \mathbf{0}_{1 \times n(x)} & \mathbf{0}_{1 \times n(z)} \end{bmatrix} = \begin{bmatrix} \mathbf{h}^y_x & \mathbf{h}^y_z \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{h}^x_x & \mathbf{h}^x_z \\ \mathbf{0}_{n(z) \times n(x)} & \Pi \end{bmatrix}}_{=:\underline{\mathbf{u}}_w}. \tag{A.13}$$

In order to evaluate the Hessian, we first apply (A.7) to the composite function $(\mathbf{h}^y \circ \mathbf{u})(\mathbf{w}_t, \sigma)$ in the thrid line of (A.8b). This yields

$$\begin{aligned}(\mathbf{h}^y \circ \mathbf{u})_{ww} &= \left(I_{n(y)} \otimes \mathbf{u}_w^T\right) \mathbf{h}^y_{ww} \mathbf{u}_w + \left(\mathbf{h}^y_{(w,\sigma)} \otimes I_{n(w)}\right) \mathbf{u}_{ww} \\ &= \left(I_{n(y)} \otimes \underline{\mathbf{u}}_w^T\right) \mathbf{h}^y_{ww} \underline{\mathbf{u}}_w + \left(\mathbf{h}^y_w \otimes I_{n(w)}\right) \underline{\mathbf{u}}_{ww},\end{aligned} \tag{A.14}$$

where the second equality makes use of (A.10). The matrix $\underline{\mathbf{u}}_w$ is given by the right-most matrix in equation (A.13). Differentiating this matrix with respect to $\mathbf{w}_t$ yields

$$\underline{\mathbf{u}}_{ww} = \begin{bmatrix} \mathbf{h}^x_{ww} \\ \mathbf{0}_{n(z)n(w) \times n(w)} \end{bmatrix} \tag{A.15}$$

so that

$$\begin{aligned}(\mathbf{h}^y \circ \mathbf{u})_{ww} &= \left(I_{n(y)} \otimes \underline{\mathbf{u}}_w^T\right) \mathbf{h}^y_{ww} \underline{\mathbf{u}}_w + \left(\mathbf{h}^y_w \otimes I_{n(w)}\right) \begin{bmatrix} \mathbf{h}^x_{ww} \\ \mathbf{0}_{n(z)n(w) \times n(w)} \end{bmatrix}, \\ &= \left(I_{n(y)} \otimes \underline{\mathbf{u}}_w^T\right) \mathbf{h}^y_{ww} \underline{\mathbf{u}}_w + \left(\mathbf{h}^y_x \otimes I_{n(w)} \quad \mathbf{h}^y_z \otimes I_{n(w)}\right) \begin{bmatrix} \mathbf{h}^x_{ww} \\ \mathbf{0}_{n(z)n(w) \times n(w)} \end{bmatrix}, \\ &= \left(I_{n(y)} \otimes \underline{\mathbf{u}}_w^T\right) \mathbf{h}^y_{ww} \underline{\mathbf{u}}_w + \left(\mathbf{h}^y_{\mathbf{x}} \otimes I_{n(w)}\right) \mathbf{h}^x_{ww}.\end{aligned} \tag{A.16}$$

The remaining parts of the Hessian of $\mathbf{f}$ are easy to compute directly from the definition in (A.8b).. Just note that, e.g., the definition $x_{it} := f^i(\mathbf{w}_t, \sigma)$, $i = 1, \ldots, n(x)$ has the Hessian matrix $f^i_{ww} = \mathbf{0}_{n(w) \times n(w)}$. Therefore, the matrix $\mathbf{f}_{ww}$ is given by

$$\mathbf{f}_{ww} = \begin{bmatrix} \mathbf{h}^x_{ww} \\ \mathbf{0}_{n(z)n(w) \times n(w)} \\ \left(I_{n(y)} \otimes \underline{\mathbf{u}}_w^T\right) \mathbf{h}^y_{ww} \underline{\mathbf{u}}_w + \left(\mathbf{h}^y_{\mathbf{x}} \otimes I_{n(w)}\right) \mathbf{h}^x_{ww} \\ \mathbf{0}_{n(x)n(w) \times n(w)} \\ \mathbf{0}_{n(z)n(w) \times n(w)} \\ \mathbf{h}^y_{ww} \end{bmatrix}. \tag{A.17}$$

In order to evaluate the second term on the right-hand side of equation (A.11) we partition the Jacobian of $\mathbf{g}$. Let $\mathbf{g}_i$, $i \in \{x, z, y, x', z', y'\}$ denote the derivatives of $\mathbf{g}$ with respect to $\mathbf{x}_t$, $\mathbf{z}_t$, $\mathbf{y}_t$, $\mathbf{x}_{t+1}$, $\mathbf{z}_{t+1}$, and $\mathbf{y}_{t+1}$. Therefore:

$$\begin{aligned}&\left(\mathbf{g}_s \otimes I_{n(w)}\right) \mathbf{f}_{ww} \\ &= \left(\mathbf{g}_{x'} \otimes I_{n(w)} \quad \mathbf{g}_{z'} \otimes I_{n(w)} \quad \mathbf{g}_{y'} \otimes I_{n(w)} \quad \mathbf{g}_x \otimes I_{n(w)} \quad \mathbf{g}_z \otimes I_{n(w)} \quad \mathbf{g}_y \otimes I_{n(w)}\right) \mathbf{f}_{ww}, \\ &= \left(\mathbf{g}_{x'} \otimes I_{n(w)}\right) \mathbf{h}^x_{ww} + \left(\mathbf{g}_{y'} \otimes I_{n(w)}\right) C_1 \mathbf{h}^y_{ww} C_2 + \left(\mathbf{g}_{y'} \otimes I_{n(w)}\right) C_3 \mathbf{h}^x_{ww} + \left(\mathbf{g}_y \otimes I_{n(w)}\right) \mathbf{h}^y_{ww}, \\ &= B_1 \mathbf{h}^y_{ww} + B_2 \mathbf{h}^x_{ww} + B_3 C_1 \mathbf{h}^y_{ww} C_2 + B_3 C_3 \mathbf{h}^x_{ww}.\end{aligned}$$

where

$$
\begin{aligned}
B_1 &= \mathbf{g}_y \otimes I_{n(w)}, \\
B_2 &= \mathbf{g}_{x'} \otimes I_{n(w)}, \\
B_3 &= \mathbf{g}_{y'} \otimes I_{n(w)}, \\
C_1 &= I_{n(y)} \otimes C_2^T, \\
C_2 &= \begin{bmatrix} \mathbf{h}_x^x & \mathbf{h}_z^x \\ \mathbf{0}_{n(z)\times n(x)} & \Pi \end{bmatrix}, \\
C_3 &= \mathbf{h}_{\mathbf{x}}^{\mathbf{y}} \otimes I_{n(w)}.
\end{aligned} \tag{A.18}
$$

Finally, let

$$
\begin{aligned}
A_1 &:= \left(I_{n(x)+n(y)} \otimes A_2^T\right) \mathbf{g}_{ss} A_2, \\
A_2 &:= \mathbf{f}_w = \begin{bmatrix} \mathbf{h}_x^x & \mathbf{h}_z^x \\ \mathbf{0}_{n(z)\times n(x)} & \Pi \\ \mathbf{h}_x^y \mathbf{h}_x^x & \mathbf{h}_x^y \mathbf{h}_z^x + \mathbf{h}_z^y \Pi \\ I_{n(x)} & \mathbf{0}_{n(x)\times n(z)} \\ \mathbf{0}_{n(z)\times n(x)} & I_{n(z)} \\ \mathbf{h}_x^y & \mathbf{h}_z^y \end{bmatrix}
\end{aligned} \tag{A.19}
$$

so that the system of equations (A.11) can be written as

$$
-A_1 = B_1 \mathbf{h}_{ww}^y + B_3 C_1 \mathbf{h}_{ww}^y C_2 + (B_2 + B_3 C_3)\, \mathbf{h}_{ww}^x. \tag{A.20}
$$

This is a linear system in the unknown coefficient matrices $\mathbf{h}_{ww}^y$ and $\mathbf{h}_{ww}^x$. A straightforward way to solve this system is to employ the vec operator:[14]

$$
\begin{aligned}
-\mathrm{vec}(A_1) &= \left(I_{n(w)} \otimes B_1\right) \mathrm{vec}(\mathbf{h}_{ww}^y) + \left(C_2^T \otimes B_3 C_1\right) \mathrm{vec}(\mathbf{h}_{ww}^y) \\
&\quad + \left(I_{n(w)} \otimes (B_2 + B_3 C_3)\right) \mathrm{vec}(\mathbf{h}_{ww}^x), \\
&= \left[I_{n(w)} \otimes B_1 + C_2^T \otimes B_3 C_1, \quad I_{n(w)} \otimes (B_2 + B_3 C_3)\right] \begin{bmatrix} \mathrm{vec}(\mathbf{h}_{ww}^y) \\ \mathrm{vec}(\mathbf{h}_{ww}^x) \end{bmatrix}.
\end{aligned} \tag{A.21}
$$

The second way to solve equation (A.20) is to note that this a generalized Sylvester equation (see Kågström and Poromaa (1994), equation (1.1))

$$
AR - LB = C, \tag{A.22a}
$$

$$
DR - LE = F. \tag{A.22b}
$$

The linear algebra package (LAPACK), whose routines are freely available on the world wide web at http://www.netlib.org/lapack/, provides routines to solve this kind of

[14]We use

$$
\mathrm{vec}(ABC) = (C^T \otimes A)\,\mathrm{vec}(B).
$$

See, e.g, Sydsæter et al. (1999), equation (23.18).

equation. To put (A.20) into the form (A.22), note first that (A.20) can be written as

$$\underbrace{-A_1}_{=:C} = \underbrace{\begin{bmatrix} B_1, & B_2 + B_3 C_3 \end{bmatrix}}_{=:A} \underbrace{\begin{bmatrix} \mathbf{h}^y_{ww} \\ \mathbf{h}^x_{ww} \end{bmatrix}}_{=:R} - \underbrace{\begin{bmatrix} B_3 C_1, \mathbf{0}_{(n(x)+n(y))n(w)\times n(x)n(w)} \end{bmatrix} \begin{bmatrix} \mathbf{h}^y_{ww} \\ \mathbf{h}^x_{ww} \end{bmatrix}}_{=:L} \underbrace{(-C_2)}_{=:B} \tag{A.23a}$$

so that (A.22b) reduces to the definition of the matrix L:

$$\underbrace{\begin{bmatrix} B_3 C_1, \mathbf{0}_{(n(x)+n(y))n(w)\times n(x)n(w)} \end{bmatrix}}_{=:D} R - L \underbrace{I_{n(w)}}_{=:E} = \underbrace{\mathbf{0}_{(n(x)+n(y))n(w)\times n(w)}}_{=:F}. \tag{A.23b}$$

Coefficients of the Perturbation Parameter. We will now employ (A.7) to the composite function $\mathbb{E}_t(\mathbf{g}\circ\mathbf{f})$ with respect to σ. The respective Hessian $(\mathbf{g}\circ\mathbf{f})_{\sigma\sigma}$ is a column vector with $n(x)+n(y)$ elements:

$$\mathbf{0}_{(n(x)+n(y))\times 1} = \mathbb{E}_t\,(\mathbf{g}\circ\mathbf{f})_{\sigma\sigma} = \mathbb{E}_t\left\{\left(I_{n(x)+n(y)} \otimes \mathbf{f}_\sigma^T\right)\mathbf{g}_{ss}\mathbf{f}_\sigma + (\mathbf{g}_s \otimes I_1)\,\mathbf{f}_{\sigma\sigma}\right\}. \tag{A.24}$$

The Jacobian of $\mathbf{f}$ with respect to σ follows from differentiating (A.8b). This gives:

$$\mathbf{f}_\sigma = \begin{bmatrix} \mathbf{h}^x_\sigma \\ \boldsymbol{\mu}_\sigma(\sigma) + \Omega\boldsymbol{\nu}_{t+1} \\ \mathbf{h}^y_u \mathbf{u}_\sigma \\ \mathbf{0}_{n(x)\times 1} \\ \mathbf{0}_{n(z)\times 1} \\ \mathbf{h}^y_\sigma \end{bmatrix}, \quad \mathbf{h}^y_u \mathbf{u}_\sigma = \begin{bmatrix} \mathbf{h}^y_x & \mathbf{h}^y_z & \mathbf{h}^y_\sigma \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{h}^x_\sigma \\ \boldsymbol{\mu}_\sigma(\sigma) + \Omega\boldsymbol{\nu}_{t+1} \\ 1 \end{bmatrix}}_{=:\mathbf{u}_\sigma}. \tag{A.25}$$

Remember that all derivatives in this expression must be evaluated at $\sigma = 0$ and $\mathbf{s} = (\mathbf{x}, \mathbf{0}_{n(z)\times 1}, \mathbf{y}, \mathbf{x}, \mathbf{0}_{n(z)\times 1}, \mathbf{y})$. It is well-known from Schmitt-Grohé and Uribe (2004b) that

$$\mathbf{h}^x_\sigma = \mathbf{0}_{n(x)\times 1}, \quad \mathbf{h}^y_\sigma = \mathbf{0}_{n(y)\times 1}, \quad \mathbf{h}^x_{w\sigma} = \mathbf{0}_{n(x)\times n(x)}, \quad \mathbf{h}^y_{w\sigma} = \mathbf{0}_{n(y)\times n(x)}. \tag{A.26}$$

These results continue to hold due to assumptions (4d) and (4e). Therefore, $\mathbf{f}_\sigma$ is equal to:

$$\mathbf{f}_\sigma = \begin{bmatrix} \mathbf{0}_{n(x)\times 1} \\ \Omega\boldsymbol{\nu}_{t+1} \\ \mathbf{h}^y_z \Omega\boldsymbol{\nu}_{t+1} \\ \mathbf{0}_{n(x)\times 1} \\ \mathbf{0}_{n(z)\times 1} \\ \mathbf{0}_{n(y)\times 1} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n(x)\times n(z)} \\ I_{n(z)} \\ \mathbf{h}^y_z \\ \mathbf{0}_{n(x)\times n(z)} \\ \mathbf{0}_{n(z)\times n(z)} \\ \mathbf{0}_{n(y)\times n(z)} \end{bmatrix} \Omega\boldsymbol{\nu}_{t+1} =: N\Omega\boldsymbol{\nu}_{t+1}. \tag{A.27}$$

Differentiating equation (A.25) again with respect to σ yields

$$\mathbf{f}_{\sigma\sigma} = \begin{bmatrix} \mathbf{h}^x_{\sigma\sigma} \\ \boldsymbol{\mu}_{\sigma\sigma} \\ M \\ \mathbf{0}_{n(x)\times 1} \\ \mathbf{0}_{n(z)\times 1} \\ \mathbf{h}^y_{\sigma\sigma} \end{bmatrix}. \tag{A.28}$$

The matrix M derives from applying (A.7) to the composite function $(\mathbf{h}^y \circ \mathbf{u})(\sigma)$.

$$M := (\mathbf{h}^y \circ \mathbf{u})_{\sigma\sigma} = M_1 + M_2 = \underbrace{\left(I_{n(y)} \otimes \mathbf{u}_\sigma^T\right) \mathbf{h}_{uu}^y \mathbf{u}_\sigma}_{=:M_1} + \underbrace{\left(\mathbf{h}_u^y \otimes I_1\right) \mathbf{u}_{\sigma\sigma}}_{=:M_2},$$

$$\mathbf{h}_{uu}^y = \begin{bmatrix} \mathbf{h}_{uu}^{y_1} \\ \mathbf{h}_{uu}^{y_2} \\ \vdots \\ \mathbf{h}_{uu}^{y_{n(y)}} \end{bmatrix}, \quad \mathbf{h}_{uu}^{y_i} = \begin{bmatrix} \mathbf{h}_{xx}^{y_i} & \mathbf{h}_{xz}^{y_i} & \mathbf{h}_{x\sigma}^{y_i} \\ \mathbf{h}_{zx}^{y_i} & \mathbf{h}_{zz}^{y_i} & \mathbf{h}_{z\sigma}^{y_i} \\ \mathbf{h}_{\sigma x}^{y_i} & \mathbf{h}_{\sigma z}^{y_i} & \mathbf{h}_{\sigma\sigma}^{y_1} \end{bmatrix}, \quad \mathbf{u}_{\sigma\sigma} = \begin{bmatrix} \mathbf{h}_{\sigma\sigma}^x \\ \boldsymbol{\mu}_{\sigma\sigma} \\ 0 \end{bmatrix}. \tag{A.29}$$

To evaluate this expression at $\sigma = 0$ and $\mathbf{s}$, we again employ (A.26) and (4e) and get

$$\begin{aligned} M_1 &= \begin{bmatrix} \mathbf{u}_\sigma^T & \mathbf{0}_{1\times(n(x)+n(z)+1)} & \cdots & \mathbf{0}_{1\times(n(x)+n(z)+1)} \\ \mathbf{0}_{1\times(n(x)+n(z)+1)} & \mathbf{u}_\sigma^T & \cdots & \mathbf{0}_{1\times(n(x)+n(z)+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1\times(n(x)+n(z)+1)} & \mathbf{0}_{1\times(n(x)+n(z)+1)} & \cdots & \mathbf{u}_\sigma^T \end{bmatrix} \begin{bmatrix} \mathbf{h}_{uu}^{y_1} \\ \mathbf{h}_{uu}^{y_2} \\ \vdots \\ \mathbf{h}_{uu}^{y_{n(y)}} \end{bmatrix} \begin{bmatrix} \mathbf{0}_{n(x)\times 1} \\ \Omega \boldsymbol{\nu}_{t+1} \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\nu}_{t+1}^T \Omega^T \mathbf{h}_{zz}^{y_1} \Omega \boldsymbol{\nu}_{t+1} \\ \boldsymbol{\nu}_{t+1}^T \Omega^T \mathbf{h}_{zz}^{y_2} \Omega \boldsymbol{\nu}_{t+1} \\ \vdots \\ \boldsymbol{\nu}_{t+1}^T \Omega^T \mathbf{h}_{zz}^{y_{n(y)}} \Omega \boldsymbol{\nu}_{t+1} \end{bmatrix} + \mathbf{h}_{\sigma\sigma}^y =: M_{11} + \mathbf{h}_{\sigma\sigma}^y \end{aligned} \tag{A.30}$$

and

$$M_2 = \begin{bmatrix} \mathbf{h}_x^y & \mathbf{h}_z^y & \mathbf{0}_{n(y)\times 1} \end{bmatrix} \begin{bmatrix} \mathbf{h}_{\sigma\sigma}^x \\ \boldsymbol{\mu}_{\sigma\sigma} \\ 0 \end{bmatrix} = \mathbf{h}_x^y \mathbf{h}_{\sigma\sigma}^x + \mathbf{h}_z^y \boldsymbol{\mu}_{\sigma\sigma}. \tag{A.31}$$

Summarizing the results in (A.27), (A.30), and (A.31), the expression in (A.24) is equal to

$$\begin{aligned} \mathbf{0}_{n(x)+n(y)} &= \mathbb{E}_t \left\{ \begin{bmatrix} (N\Omega\boldsymbol{\nu}_{t+1})^T g_{ss}^1 (N\Omega\boldsymbol{\nu}_{t+1}) \\ (N\Omega\boldsymbol{\nu}_{t+1})^T g_{ss}^2 (N\Omega\boldsymbol{\nu}_{t+1}) \\ \vdots \\ (N\Omega\boldsymbol{\nu}_{t+1})^T g_{ss}^{n(x)+n(y)} (N\Omega\boldsymbol{\nu}_{t+1}) \end{bmatrix} \right\} \\ &\quad + \mathbb{E}_t \left\{ \begin{bmatrix} \mathbf{g}_{x'} & \mathbf{g}_{z'} & \mathbf{g}_{y'} & \mathbf{g}_x & \mathbf{g}_z & \mathbf{g}_y \end{bmatrix} \begin{bmatrix} \mathbf{h}_{\sigma\sigma}^x \\ \boldsymbol{\mu}_{\sigma\sigma} \\ M_1 + M_2 \\ \mathbf{0}_{n(x)\times 1} \\ \mathbf{0}_{n(z)\times 1} \\ \mathbf{h}_{\sigma\sigma}^y \end{bmatrix} \right\} \\ &= \mathbb{E}_t \left\{ \begin{bmatrix} (N\Omega\boldsymbol{\nu}_{t+1})^T g_{ss}^1 (N\Omega\boldsymbol{\nu}_{t+1}) \\ (N\Omega\boldsymbol{\nu}_{t+1})^T g_{ss}^2 (N\Omega\boldsymbol{\nu}_{t+1}) \\ \vdots \\ (N\Omega\boldsymbol{\nu}_{t+1})^T g_{ss}^{n(x)+n(y)} (N\Omega\boldsymbol{\nu}_{t+1}) \end{bmatrix} \right\} \\ &\quad + \mathbf{g}_{x'} \mathbf{h}_{\sigma\sigma}^y + \mathbf{g}_{z'} \boldsymbol{\mu}_{\sigma\sigma} + \mathbf{g}_{y'} \mathbb{E}_t M_{11} + \mathbf{g}_{y'} \mathbf{h}_{\sigma\sigma}^y + \mathbf{g}_{y'} \mathbf{h}_x^y \mathbf{h}_{\sigma\sigma}^x + + \mathbf{g}_{y'} \mathbf{h}_z^y \boldsymbol{\mu}_{\sigma\sigma} + \mathbf{g}_y \mathbf{h}_{\sigma\sigma}^y. \end{aligned}$$

The final step is to evaluate the conditional expectations in this expression. First note that each element

$$\Delta_i = (N\Omega \boldsymbol{\nu}_{t+1})^T g^i_{ss} (N\Omega \boldsymbol{\nu}_{t+1}) \quad = \underbrace{\boldsymbol{\nu}^T_{t+1}\Omega^T}_{\Delta_{i1}} \underbrace{N^T g^i_{ss} (N\Omega \boldsymbol{\nu}_{t+1})}_{=:\Delta_{i2}}$$

of the first term in the previous equation is a scalar $\Delta_i = \Delta_{i1}\Delta_{i2}$. Per definition, the trace of a scalar is equal to the scalar itself, $\mathrm{tr}(\Delta_i) = \Delta_i$ so that we can apply a well known result of the trace operator[15]

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Thus,

$$\begin{aligned}\mathbb{E}_t\Delta_i &= \mathbb{E}_t \,\mathrm{tr}(\Delta_i) = \mathbb{E}_t \,\mathrm{tr}(\Delta_{i1}\Delta_{i2}) = \mathbb{E}_t \,\mathrm{tr}(\Delta_{i2}\Delta_{i1}) = \mathrm{tr}(\mathbb{E}_t(\Delta_{i2}\Delta_{i1})) \\ &= \mathrm{tr}\left(N^T g^i_{ss} N \mathbb{E}_t \left(\Omega \boldsymbol{\nu}_{t+1} \boldsymbol{\nu}^T_{t+1} \boldsymbol{\Omega}^T\right)\right) \overset{(4c)}{=} \mathrm{tr}\left(N^T g^i_{ss} N\Omega\Omega^T\right).\end{aligned}$$

In the same way we can determine $\mathbb{E}_t M_{11}$:

$$\mathbb{E}_t M_{11} = \begin{bmatrix} \mathrm{tr}\left(\mathbf{h}^{y_1}_{zz}\Omega\Omega^T\right) \\ \mathrm{tr}\left(\mathbf{h}^{y_2}_{zz}\Omega\Omega^T\right) \\ \vdots \\ \mathrm{tr}\left(\mathbf{h}^{y_{n(y)}}_{zz}\Omega\Omega^T\right) \end{bmatrix}.$$

Finally, consider an extension of the trace operator proposed by Gomme and Klein (2011):

$$\mathrm{trm}\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix} := \begin{bmatrix} \mathrm{tr}(X_1) \\ \mathrm{tr}(X_2) \\ \vdots \\ \mathrm{tr}(X_m) \end{bmatrix}. \tag{A.32}$$

This operator returns the traces of the m blocks of the $mn \times n$ matrix X in the m-dimensional column vector $\mathbf{x}$. Accordingly, the system of linear equations (A.24) can be written as:

$$\begin{aligned}&-\mathrm{trm}\left(\left(I_{n(x)+n(y)} \otimes N^T\right) \mathbf{g}_{ss} N\Omega\Omega^T\right) - \mathbf{g}_{y'}\,\mathrm{trm}\left(\left(I_{n(y)} \otimes (\Omega\Omega^T)\right) \mathbf{h}^y_{zz}\right) \\ &-\left(\mathbf{g}_{z'} + \mathbf{g}_{y'}\mathbf{h}^y_z\right) \boldsymbol{\mu}_{\sigma\sigma} \\ &= \left[\mathbf{g}_y + \mathbf{g}_{y'} \quad \mathbf{g}_{x'} + \mathbf{g}_{y'}\mathbf{h}^y_x\right] \begin{bmatrix} \mathbf{h}^y_{\sigma\sigma} \\ \mathbf{h}^x_{\sigma\sigma} \end{bmatrix}.\end{aligned} \tag{A.33}$$

The term in the second line of this expression distinguishes the standard solution from our extended version of the canonical DSGE model in (4).

[15] See, e.g., Sydsæter et al. (1999), equation (19.22).

C Weighted Residuals Methods

C.1 Residual Functions

Consider the equilibrium conditions of the model in (A.1a) and let

$$h: \mathbb{R}_+ \times \mathbb{R} \to [0,1] \quad \text{and} \quad V: \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$$

with $z_t := \ln A_t$ denote the policy function for n_t and the value function V_t, respectively:

$$n_t = h(k_t, z_t) \quad \text{and} \quad V_t = V(k_t, z_t).$$

First, observe that given the state variables k_t, z_t and the policy function h we can successively solve for the remaining variables as follows:

$$y_t = y(k_t, z_t, h) \overset{\text{(A.1a)}}{:=} \exp(z_t) k_t^{\theta} h(k_t, z_t)^{1-\theta}, \tag{A.34}$$

$$c_t = c(k_t, z_t, h) \overset{\text{(A.1c)}}{:=} (1-\theta) \frac{\alpha}{1-\alpha} \frac{y(k_t, z_t, h)}{h(k_t, z_t))} (1 - h(k_t, z_t)), \tag{A.35}$$

$$\lambda_t = \lambda(k_t, z_t, h) \overset{\text{(A.1b)}}{:=} \alpha c(k_t, z_t, h)^{\alpha(1-\sigma)-1} (1 - h(k_t, z_t))^{(1-\alpha)(1-\sigma)}, \tag{A.36}$$

$$k_{t+1} = k(k_t, z_t, h) \overset{\text{(A.1e)}}{:=} (1-\delta) k_t + y(k_t, z_t, h) - c(k_t, z_t, h), \tag{A.37}$$

$$z_{t+1} = \rho z_t + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N\left(-\frac{\tau^2}{2(1+\rho)}, \tau^2\right). \tag{A.38}$$

Therefore all period t variables can be expressed by k_t, z_t and $h(k_t, z_t)$ while all variables of period $t+1$ can be expressed by $k_t, z_t, h(k_t, z_t)$ and ϵ_{t+1}. We introduce the notation

$$\begin{aligned}\text{rhs}_1(k_t, z_t, \epsilon_{t+1}, h) &:= \beta \frac{\lambda_{t+1}}{\lambda_t} \left(\theta \frac{y_{t+1}}{k_{t+1}} + 1 - \delta\right) = \\ &= \beta \frac{\lambda(k(k_t, z_t, h), \rho z_t + \epsilon_{t+1}, h)}{\lambda(k_t, z_t, h)} \left(\theta \frac{y(k(k_t, z_t, h), \rho z_t + \epsilon_{t+1}, h)}{k(k_t, z_t, h)} + 1 - \delta\right)\end{aligned} \tag{A.39a}$$

and

$$\text{rhs}_2(k_t, z_t, \epsilon_{t+1}, h, V) := V(k(k_t, z_t, h), \rho z_t + \epsilon_{t+1}) \tag{A.39b}$$

From (A.1f) the policy function h solves the following equation for all k_t and z_t

$$R_1(k_t, z_t, h) := 1 - \int_{\mathbb{R}} \text{rhs}_1(k_t, z_t, \epsilon_{t+1}, h) \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{\left(\epsilon_{t+1} + \frac{\tau^2}{2(1+\rho)}\right)^2}{2\tau^2}\right) d\epsilon_{t+1} = 0. \tag{A.40}$$

Analogously from (A.1g), V satisfies

$$\begin{aligned}R_2(k_t, z_t, h, V) &:= V(k_t, z_t) - \frac{c(k_t, z_t, h)^{\alpha(1-\eta)} (1 - h(k_t, z_t))^{(1-\alpha)(1-\eta)}}{1-\eta} \\ &- \beta \int_{\mathbb{R}} \text{rhs}_2(k_t, z_t, \epsilon_{t+1}, h, V) \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{\left(\epsilon_{t+1} + \frac{\tau^2}{2(1+\rho)}\right)^2}{2\tau^2}\right) d\epsilon_{t+1} = 0\end{aligned} \tag{A.41a}$$

for all k_t and z_t if $\eta \neq 0$ or

$$R_2(k_t, z_t, h, V) := V(k_t, z_t) - \alpha \ln(c(k_t, z_t, h)) - (1-\alpha)\ln(1 - h(k_t, z_t))$$
$$- \beta \int_{\mathbb{R}} \text{rhs}_2(k_t, z_t, \epsilon_{t+1}, h, V) \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{\left(\epsilon_{t+1} + \frac{\tau^2}{2(1+\rho)}\right)^2}{2\tau^2} \right) d\epsilon_{t+1} = 0 \tag{A.41b}$$

for all k_t and z_t if $\eta = 1$.

Suppose we have approximated h by some function $\hat{h}$ and V by $\hat{V}$. Replacing h and V in the residual functions R_1 and R_2 yields functions $R_1(k_t, z_t, \hat{h})$ and $R_2(k_t, z_t, \hat{h}, \hat{V})$. In order to evaluate these functions at a given point (k_t, z_t) we have to compute the respective integrals. We use Gauss-Hermite quadrature with M points e_i and weights w_i, $i = 1, 2, \ldots, M$ to perform this task. Thus, we define the residual functions

$$\hat{R}_1(k_t, z_t, \hat{h}) := 1 - \frac{1}{\sqrt{\pi}} \sum_{i=1}^{M} w_i \text{rhs}_1\left(k_t, z_t, \sqrt{2}\tau e_i - \frac{\tau^2}{2(1+\rho)}, \hat{h} \right), \tag{A.42a}$$

$$\hat{R}_2(k_t, z_t, \hat{h}, \hat{V}) := \hat{V}(k_t, z_t) - \begin{cases} \frac{1}{1-\eta}\left[c(k_t, z_t, \hat{h})^{\alpha(1-\eta)} (1 - \hat{h}(k_t, z_t))^{(1-\alpha)(1-\eta)} \right] & \text{for } \eta \neq 1 \\ \alpha \ln(c(k_t, z_t, \hat{h})) + (1-\alpha)\ln(1 - \hat{h}(k_t, z_t)) & \text{for } \eta = 1 \end{cases}$$
$$- \frac{\beta}{\sqrt{\pi}} \sum_{i=1}^{M} w_i \text{rhs}_2\left(k_t, z_t, \sqrt{2}\tau e_i - \frac{\tau^2}{2(1+\rho)}, \hat{h}, \hat{V} \right). \tag{A.42b}$$

C.2 Finite Element Method

We first choose a discrete grid $\Gamma \subset \mathbb{R}^2$ for the state variables (k_t, z_t) with a total of N grid points. With $S_{\Gamma,\mathbf{v}} : \mathbb{R}^2 \to \mathbb{R}$ defined as the cubic $\mathscr{C}^2$ spline interpolating the function values in $\mathbf{v} \in \mathbb{R}^N$ over the grid Γ, we search for appropriate approximations of h and V within the class

$$\left\{ \hat{h}_{\mathbf{v}_1} = S_{\Gamma,\mathbf{v}_1} : \mathbb{R}^2 \to \mathbb{R} | \mathbf{v}_1 \in \mathbb{R}^N \right\} \quad \text{and} \quad \left\{ \hat{V}_{\mathbf{v}_2} = S_{\Gamma,\mathbf{v}_2} : \mathbb{R}^2 \to \mathbb{R} | \mathbf{v}_2 \in \mathbb{R}^N \right\}.$$

Next, we determine the unknown function values in $\mathbf{v}_1$ and $\mathbf{v}_2$ at the grid points by imposing the conditions that the residuals vanish at all grid points in Γ. I.e. we solve the following system of equations for $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$\left. \begin{array}{l} 0 = \hat{R}_1(k_t, z_t, \hat{h}_{\mathbf{v}_1}), \\ 0 = \hat{R}_2(k_t, z_t, \hat{h}_{\mathbf{v}_1}, \hat{V}_{\mathbf{v}_2}) \end{array} \right\} \forall (k_t, z_t) \in \Gamma. \tag{A.43}$$

C.3 Chebyshev-Galerkin Method

We approximate the functions $h(k_t, z_t)$ and $V(k_t, z_t)$ by combinations of Chebyshev polynomials, i.e.:

$$\hat{h}(k, z) := \sum_{i=1}^{d_1^n} \sum_{j=1}^{d_2^n} \phi_{ij}^n T_{i-1}(\psi_1(k)) T_{j-1}(\psi_2(z)), \tag{A.44a}$$

$$\hat{V}(k,z) := \sum_{i=1}^{d_1^V}\sum_{j=1}^{d_2^V}\phi_{ij}^V T_{i-1}(\psi_1(k))T_{j-1}(\psi_2(z)). \tag{A.44b}$$

Since the domain of the Chebyshev polynomials of order $l = 0,1,2,\ldots$, $T_l(x)$ is the interval $x \in [-1,1]$, we must choose boundaries for the stock of capital k and the log of the TFP shock z and map the respective intervals $[\underline{k},\bar{k}]$ and $[\underline{z},\bar{z}]$ to this domain. The functions ψ_j, $j = 1,2$ perform this task. Let r_t, $t = 1,2,\ldots,m_1$ and q_s, $s = 1,2,\ldots,m_2$ denote the roots of the Chebyshev polynomials of order m_1 and m_2, respectively. Furthermore, let $\tilde{\psi}_1$ and $\tilde{\psi}_2$ define the inverse mappings from $[-1,1]$ to $[\underline{k},\bar{k}]$ and $[\underline{z},\bar{z}]$, respectively. We determine the coefficients from (A.44) as solutions the system of $d_1 d_2$ non-linear equations given by:

$$0 = \frac{\pi}{m_1}\frac{\pi}{m_2}\sum_{t=1}^{m_1}\sum_{s=1}^{m_2}\hat{R}_1(\tilde{\psi}_1(r_t)),\tilde{\psi}_2(q_s),\hat{h})T_{i-1}(r_t)T_{j-1}(q_s) \quad \forall i = 1,\ldots d_1^n, j = 1,\ldots,d_2^n), \tag{A.45a}$$

$$0 = \frac{\pi}{m_1}\frac{\pi}{m_2}\sum_{t=1}^{m_1}\sum_{s=1}^{m_2}\hat{R}_2(\tilde{\psi}_1(r_t)),\tilde{\psi}_2(q_s),\hat{h},\hat{V})T_{i-1}(r_t)T_{j-1}(q_s)\; \forall i = 1,\ldots,d_1^V, j = 1,\ldots,d_2^V. \tag{A.45b}$$