mea – Mannheimer Forschungsinstitut Ökonomie und Demographischer Wandel
Gebäude L 13, 17_D-68131 Mannheim_Sekretariat +49 621 181-2773/1862_Telefax +49 621 181-1863_www.mea.uni-mannheim.de

Improving Tatonnement Methods for Solving Heterogeneous Agent Models

Alexander Ludwig

58-2004

September 2004

Improving Tatonnement Methods for Solving Heterogeneous Agent Models*

Alexander Ludwig†
MEA, University of Mannheim

This Version: September 7, 2004

Abstract

This paper modifies standard block Gauss-Seidel iterations used by tatonnement methods for solving large scale deterministic heterogeneous agent models. The composite method between *first-* and *second-order tatonnement* methods is shown to considerably improve convergence both in terms of speed as well as robustness relative to conventional first-order tatonnement methods. In addition, the relative advantage of the modified algorithm increases in the size and complexity of the economic model. Therefore, the algorithm allows significant reductions in computational time when solving large models. The algorithm is particularly attractive since it is easy to implement - it only augments conventional and intuitive tatonnement iterations with standard numerical methods.

JEC classification: C63, C68, E13
Keywords: OLG models; Gauss-Seidel iterations; Quasi-Newton methods

*I thank Alan Auerbach, Axel Börsch-Supan, Wouter Den Haan, Ken Judd, Michel Juillard, Gabriele Steidl and Joachim Winter as well as several seminar participants at the University of Mannheim, at the 2004 SED Annual Meeting in Florence and the 2004 EEA Annual Meeting in Madrid for helpful comments. Financial support from the Deutsche Forschungsgemeinschaft, SFB 504, at the University of Mannheim and from the Gesamtverband der Deutschen Versicherungswirtschaft is gratefully acknowledged.

†MEA, University of Mannheim; L13,17; 68131 Mannheim; Germany. Phone: +49-621-181-1866, Fax:+49-621-181-1863, email: alexander.ludwig@mea.uni-mannheim.de

1 Introduction

This paper modifies Gauss-Seidel iterations used to solve large-scale deterministic heterogeneous agent models. Such models are increasingly used for analysis of economic questions. Standard procedures use domain truncation methods and resort to general methods for solving large systems of (nonlinear) equations. Three types of such conventional solution methods can be distinguished: (i) Newton based methods such as the L-B-J method[1], (ii) the Fair-Taylor (extended path) method[2] and (iii) tatonnement methods[3], see Judd et al. (2000). These conventional methods have in common that they solve the model for each time t element of all endogenous variables.[4]

This paper is concerned with traditional methods. Conventional *first-order tatonnement* methods are commonly used to solve large-scale overlapping generations (OLG) models and the particular application chosen in this paper is a standard OLG model in the tradition of Auerbach and Kotlikoff (1987). As a composite of first- and second-order tatonnement methods, the algorithm developed here is a straightforward modification of such conventional methods. The analysis shows that this hybrid method greatly improves convergence relative to standard first-order methods.

While L-B-J and Fair-Taylor methods regard any perfect foresight general equilibrium model simply as a system of (non-linear) equations including aggregate and disaggregate variables and iterate over this entire system, tatonnement methods break variables into *aggregate* and *disaggregate* variables. Outer loops then proceed via block Gauss-Seidel algorithms using aggregate variables only, whereas inner loops are used to solve for disaggregate variables in a (separate) disaggregate model. Outer loops work as follows: Let $P = S^{-1}(Q)$ denote a sequence of factor prices corresponding to sequences of factor supplies Q, where S^{-1} denotes the inverse supply function. Equilibrium of tatonnement methods is defined as a fixed point, $Q = D(S^{-1}(Q))$, where D denotes the demand function. $S^{-1}(Q)$ and $D(P)$ are solved by inner loops of the disaggregate model and by aggregating individual decisions. The fixed point problem suggests to execute the iteration $Q^{k+1} = D(P^{k+1}) = D(S^{-1}(Q^k))$, which is the familiar *hog-cycle* process, where k is the iteration number.[5] Depending on the functional form of S relative to D such iterations may however not converge. These convergence problems force researchers to rely on *ad hoc* dampening factors such that the

[1] See Laffargue (1990), Boucekkine (1995), Juillard (1996) and Juillard et al. (1998).

[2] See Fair and Taylor (1983).

[3] See Auerbach and Kotlikoff (1987).

[4] More recently, Judd (2002) has proposed an alternative route. Rather than explicitely solving for each time t element, Judd suggests to use prior information about the time path of the endogenous variables and to approximate it by a functional form with a low-dimensional parameter vector. Judd's method can be regarded as a more modern approach.

[5] Since P^{k+1} and not P^k is used to form an update of Q^{k+1} the iterations performed are non-linear Block-Gauss-Seidel iterations.

iteration rewrites as $Q^{k+1} = Q^k - w(Q^k - D(S^{-1}(Q^k))) = Q^k - w(Q^k - \tilde{Q}^k)$, where w is the dampening factor, the relative weight w attached to Q^k and $\tilde{Q}^k$ respectively.[6]

Such modifications of standard Gauss-Seidel iterations have also been referred to as *fast* Gauss-Seidel (FGS) iterations (Hughes Hallet 1984).[7] Since only values of $D(S^{-1}(Q^k))$ are used to solve the fixed point problem and no additional information on the functional form of D, respectively S, these methods belong to the class of first-order iterative methods. While intuitive, convergence of these methods is slow (linear at best) and they may not converge at all even after various dampening factors have been tried out. As an alternative to using *ad hoc* dampening factors, *optimal* dampening factors can be determined. However, they are difficult to determine even for linear models, see, e.g., Hagemann and Young (1981) and Judd (1999). Therefore, various adaptive techniques to update dampening factors as the iteration proceeds have been suggested in the literature (Hagemann and Young 1981; Hughes Hallet 1982).

As an alternative to such first-order iterations, *second-order tatonnnement* methods may be used. Fixed point problems such as $Q = D(S^{-1}(Q))$ can be transformed to a root-finding problem which suggests to iterate as $Q^{k+1} = Q^k - [J(Q^k)]^{-1}(Q^k - D(S^{-1}(Q^k))) = Q^k - [J(Q^k)]^{-1}G(Q^k)$ where $G(Q^k)$ is a system of simultaneous non-linear equations, Q^k is the root of these equations and $J(Q^k)$ is the Jacobi matrix. Such systems may be solved using standard non-linear equation solvers, see, e.g., Feroli (2002) and Domeij and Floden (2003) for applications in an OLG context using relatively simple models. Since the dimension of the Jacobi matrix is $mT \times mT$, second-order methods become costly as the dimension of T or m, and therefore the complexity of the economic model, increases.

Against this background, this paper suggests to use a composite of standard first-order iterations and second-order methods by combining Gauss-Seidel iterations with Quasi-Newton methods[8]. The algorithm will therefore by referred to as Gauss-Seidel-Quasi-Newton method (GSQN). By economic insight the dimension of the Jacobi matrix is reduced for the system $G(Q^k)$ of non-linear equations characterizing steady state situations. Since certain transformations of economic variables in Q (and P) are constant in the steady state of economic models, the exact Jacobi matrix is shown to be given by $J = W^{-1} \otimes I$ where W is of dimension $m \times m$. Since m is generally quite small - for a standard one sector closed economy general equilibrium growth model with endogenous capital formation and endogenous labor supply m equals 2 - the Jacobi matrix can easily be determined by standard finite difference methods in fast steady state iterations. For transition iterations, the matrix is used as an approximate Jacobi matrix

[6]Note that dampening factors play a similar role as adaptive expectations in the familiar cobweb model.

[7]For convergent problems, w may also be set such as to accelerate convergence.

[8]An extensive treatment of similar methods can be found in Ortega and Rheinboldt (2000).

and updated by Broyden's method as the iterations proceeds. Accordingly, the matrix W may be interpreted as an approximate Jacobi matrix or as a matrix of multiple dampening factors (Hughes Hallet 1984). The attractiveness of GSQN stems from its simplicity: the intuitive appeal and relatively low computational demands of tatonnement iterations are combined with standard Newton based methods that are implementable at little extra cost.

As an illustration of the GSQN procedure, two economic models are developed. The first is a simple static hog-cycle model that is only used to shed light on the strong *economic* restrictions implicit in one-parameter fixed dampening. The second model is a large-scale dynamic multi-country overlapping generations (OLG) model with endogenous labor supply. It is used for simulations to compare the relative performance of the fast Gauss-Seidel algorithm (FGS) with the Gauss-Seidel-Quasi-Newton (GSQN) algorithm under various combinations of structural model parameters. In addition, the dimension m is increased from $m = 1$ (closed economy model with exogenous labor supply) to $m = 4$ (three country model with endogenous labor supply). Previewing results, the simple modifications suggested in this paper quite considerably improve convergence when compared to standard FGS. For the latter, only relatively low values of the dampening factor such as $w = 0.1$ lead to convergence for all cases considered. For higher values of w, robustness of FGS is found to decrease sharply: for $w = 0.3$, FGS does not converge for up to 40 percent of cases. In contrast, GSQN converges for all these simulations. For transition calculations, average convergence speeds of GSQN are about two times higher than those of FGS with $w = 0.1$ when $m = 1$ and about *seven* times higher when $m = 4$. Hence, GSQN considerably improves convergence both in terms of speed and in terms of robustness relative to standard FGS. The increase of the relative advantage of GSQN relative to FGS as m increases is due to the fact that the restrictions on the true Jacobi matrix of the system of equations $G(Q^k)$ imply constant (and equal) elements along the diagonal and off-diagonal elements to be equal to zero. As the dimension m increases, the loss of information implied by these restrictions becomes more and more costly. Therefore, GSQN is of particular advantage for large and therefore more complex models.

The paper proceeds as follows: Section 2 provides some general definitions and a brief review of tatonnement methods. Section 3 develops the suggested modification of the conventional Gauss-Seidel algorithm, GSQN. Section 4 contains the above mentioned economic examples used to illustrate GSQN and its differences to FGS. Section 5 compares the relative performances of FGS and GSQN for the OLG model developed in Section 4. Section 6 concludes.

2 Tatonnement Methods

Let $Y = \{y_i\}_{i=1}^{n}$ where $y_i = \{y_{i,t}\}_{t=0}^{T} \forall i$ be a list of all endogenous variables of the economic model. For example, y_i includes wage rates and interest rates as aggregate variables (a_i) as well as disaggregate variables (b_i) such as consumption and assets of individual households, etc. Note that $b_i = \{\{b_{i,e,t}\}_{e=0}^{E_i}\}_{t=0}^{T}$ where the number of disaggregate units e may differ across i. Collect $A = (a_1, a_2, ...)$ and $B = (b_1, b_2, ...)$. For further reference, split A as $A = (Q, P)$ where Q are aggregate factor supply variables such as the aggregate capital stock and aggregate labor supply of an economy and P are the associated factor price variables such as aggregate interest and wage rates and let $Q = (q_1, ..., q_m) \in \mathbb{R}^{mT}$ as well as $P = (p_1, ..., p_m) \in \mathbb{R}^{mT}$, compare Section 1. Further, let $Z = (z_1, z_2, ...)$ be a list of exogenous variables such as population data of cohorts living at time t. Note that some z_i may be disaggregate variables as well. Deterministic perfect foresight heterogeneous agent models can be written in a general form as

$$\begin{aligned} F(Y, Z) &= 0 \\ y_{i,0} = \bar{y}_{i,0},\ i = 0, 1, ..., n_i,\ n_i < n & \\ y_{i,t} \text{ bounded for all } i & \end{aligned} \tag{1}$$

where $F(Y, Z)$ are nT equations of non-linear functions that represent equilibrium. Since Z are exogenous they will be dropped from here on. The equations in (1) include Euler equations, asset accumulation equations, market clearing conditions as well as any other equations that define equilibrium. Domain truncation has been applied in equation (1) since the time horizon starts in period $t = 0$ departing from some initial conditions and is restricted to T.

Solution methods such as Fair-Taylor and L-B-J directly solve systems of equations such as (1) for each element in $y_{i,t}$ by Gauss-Seidel iterations or Newton based methods respectively. In contrast, tatonnement methods break the system of equations in (1) into a factor *supply* and a factor *demand* model. Both require inner loops to solve and to aggregate individual decision problems. A perfect foresight OLG model of the form given in equation (1) can be re-written as

$$\begin{aligned} \textit{Supply model: } P &= S^{-1}(Q) \\ \textit{Demand model: } Q &= D(P) \\ \textit{Aggregation: } P = S^{-1}(Q) &= \Sigma^s(B^s(Q)) \\ \text{and } Q = D(P) &= \Sigma^d(B^d(P)), \end{aligned} \tag{2}$$

where S^{-1} is the inverse aggregate supply function and D is the aggregate demand function. B^s (B^d) are supply (demand) side disaggregate variables and $B = (B^s, B^d)$. The aggregators, Σ^d and Σ^s, are only used to indicate that aggregate demand and supply functions are derived from individual decisions of heterogeneous agents and will be ignored from here on. Combining the first two

lines of equation (2) leads to the definition of equilibrium of a heterogeneous agent model as a fixed point given by

$$Q = D(S^{-1}(Q)), \tag{3}$$

where $Q - D(S^{-1}(Q))$ are m equations of non-linear functions that define equilibrium.

The fixed-point equation in (3) suggests to use standard (block) Gauss-Seidel iterations to solve for Q and hence to iterate over the system[9]

$$\begin{aligned} P^{k+1} &= S^{-1}(Q^k) \\ Q^{k+1} &= D(P^{k+1}). \end{aligned}$$

This simple form ignores updates of disaggregate variables. A more general case will be discussed in Section 3.4.

Equation (4) can be more concisely written as a Gauss-Seidel fixed point iteration

$$Q^{k+1} = D(S^{-1}(Q^k)). \tag{4}$$

It is well-known that such iterations may not converge. Therefore, a dampening factor may be applied. Gauss-Seidel iterations with one-parameter fixed dampening with factor $0 < w < 1$ iterate on

$$Q^{k+1} = Q^k - w\left(Q^k - D(S^{-1}(Q^k))\right), \tag{5}$$

compare Auerbach and Kotlikoff (1987, pp. 46-50). In case the fixed-point iteration in equation (4) is convergent, w may be used to accelerate convergence in which case $w > 1$. But even if such fixed-point iterations converge, convergence is slow and linear at best.

An alternative to fixed point iterations is to transform equation (3) into a root-finding problem as

$$G(Q) = Q - H(Q) = Q - D(S^{-1}(Q)) = 0,$$

where $H(\cdot)$ is introduced as a shorthand notation for $D(S^{-1})(\cdot)$.

Applying a first-order Taylor series approximation to equation (2) leads to the familiar Newton updating formula of Q given by

$$Q^{k+1} = Q^k - J^{-1}[Q^k]G(Q^k), \tag{6}$$

where $J[Q^k]$ is the Jacobi matrix of the system of equations in (2) evaluated at Q^k. Recently, several authors have used general purpose rootfinding methods

[9]The - generally less efficient - (block) Gauss-Jacobi method may be used as an alternative in which case $P^{k+1} = S^{-1}(Q^k)$ and $Q^{k+1} = D(P^k)$. Hence, rather than using P^{k+1} resulting from the first block, Gauss-Jacobi uses P^k resulting from previous iterations to form an update of Q in the second block.

to solve such problems in the OLG context, e.g., Feroli (2002) and Domeij and Floden (2003) for relatively simple models. However, as the complexity of the economic model and therefore the dimension of m and T increases, such methods become costly.

Rewriting equation (5) as

$$Q^{k+1} = Q^k - w\left(Q^k - D(S^{-1}(Q^k))\right) = Q^k - wI_{(mT\times mT)}G(Q^k)$$

then makes it obvious that Gauss-Seidel iterations with one-parameter fixed dampening restrict the elements of the true Jacobi matrix $J[Q^k]$ to $w^{-1}I_{(mT\times mT)}$. These restrictions may be summarized as follows: first, the iteration matrix is constant across all iteration steps k, second, elements along the diagonal are restricted to be equal and third, off-diagonal elements are restricted to zero. An economic interpretation of such restrictions for a stylized hog-cycle model is given below in Section 4.1.

3 The Gauss-Seidel-Quasi-Newton Method

This paper suggests an alternative to pure first- or second-order tatonnement methods by reducing the dimension of the Jacobi matrix in equation (6). For further reference and in order to highlight the restrictions implied by standard first-order methods, it will be useful to derive explicit expressions for the elements of the Jacobi matrix. Recall that $Q = \{q_i\}_{i=1}^m$, where $q_i = \{q_{i,t}\}_{t=1}^T$. Due to the specific form of the functions $G = \{g_i(Q)\}_{i=1}^m$ where $g_i(Q) = \{g_{i,t}(Q)\}_{t=0}^T$ in equation (2), the elements of the Jacobi matrix given by

$$J[Q^k] = \begin{bmatrix} \frac{\partial g_{1,0}(Q^k)}{\partial q_{1,0}^k} & \frac{\partial g_{1,0}(Q^k)}{\partial q_{1,1}^k} & \dots & \frac{\partial g_{1,0}(Q^k)}{\partial q_{1,T}^k} & \frac{\partial g_{1,0}(Q^k)}{\partial q_{2,0}^k} & \frac{\partial g_{1,0}(Q^k)}{\partial q_{2,1}^k} & \dots & \frac{\partial g_{1,0}(Q^k)}{\partial q_{2,T}^k} & \dots \\ \frac{\partial g_{1,1}(Q^k)}{\partial q_{1,0}^k} & \frac{\partial g_{1,1}(Q^k)}{\partial q_{1,1}^k} & \dots & \frac{\partial g_{1,1}(Q^k)}{\partial q_{1,T}^k} & \frac{\partial g_{1,1}(Q^k)}{\partial q_{2,0}^k} & \frac{\partial g_{1,1}(Q^k)}{\partial q_{2,1}^k} & \dots & \frac{\partial g_{1,1}(Q^k)}{\partial q_{2,T}^k} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial g_{2,0}(Q^k)}{\partial q_{1,0}^k} & \frac{\partial g_{2,0}(Q^k)}{\partial q_{1,1}^k} & \dots & \frac{\partial g_{2,0}(Q^k)}{\partial q_{1,T}^k} & \frac{\partial g_{2,0}(Q^k)}{\partial q_{2,0}^k} & \frac{\partial g_{1,0}(Q^k)}{\partial q_{2,1}^k} & \dots & \frac{\partial g_{2,0}(Q^k)}{\partial q_{2,T}^k} & \dots \\ \frac{\partial g_{2,1}(Q^k)}{\partial q_{1,0}^k} & \frac{\partial g_{2,1}(Q^k)}{\partial q_{1,1}^k} & \dots & \frac{\partial g_{2,1}(Q^k)}{\partial q_{1,T}^k} & \frac{\partial g_{2,1}(Q^k)}{\partial q_{2,0}^k} & \frac{\partial g_{1,1}(Q^k)}{\partial q_{2,1}^k} & \dots & \frac{\partial g_{2,1}(Q^k)}{\partial q_{2,T}^k} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

can be re-written as

$$\begin{bmatrix} 1 - \frac{\partial h_{1,0}(Q^k)}{\partial q_{1,0}^k} & -\frac{\partial h_{1,0}(Q^k)}{\partial q_{1,1}^k} & \dots & -\frac{\partial h_{1,0}(Q^k)}{\partial q_{1,T}^k} & -\frac{\partial h_{1,0}(Q^k)}{\partial q_{2,0}^k} & -\frac{\partial h_{1,0}(Q^k)}{\partial q_{2,1}^k} & \dots \\ -\frac{\partial h_{1,1}(Q^k)}{\partial q_{1,0}^k} & 1 - \frac{\partial h_{1,1}(Q^k)}{\partial q_{1,1}^k} & \dots & -\frac{\partial h_{1,1}(Q^k)}{\partial q_{1,T}^k} & -\frac{\partial h_{1,1}(Q^k)}{\partial q_{2,0}^k} & -\frac{\partial h_{1,1}(Q^k)}{\partial q_{2,1}^k} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -\frac{\partial h_{2,0}(Q^k)}{\partial q_{1,0}^k} & -\frac{\partial h_{2,0}(Q^k)}{\partial q_{1,1}^k} & \dots & -\frac{\partial h_{2,0}(Q^k)}{\partial q_{1,T}^k} & 1 - \frac{\partial h_{2,0}(Q^k)}{\partial q_{2,0}^k} & -\frac{\partial h_{1,0}(Q^k)}{\partial q_{2,1}^k} & \dots \\ -\frac{\partial h_{2,1}(Q^k)}{\partial q_{1,0}^k} & -\frac{\partial h_{2,1}(Q^k)}{\partial q_{1,1}^k} & \dots & -\frac{\partial h_{2,1}(Q^k)}{\partial q_{1,T}^k} & -\frac{\partial h_{2,1}(Q^k)}{\partial q_{2,0}^k} & 1 - \frac{\partial h_{1,1}(Q^k)}{\partial q_{2,1}^k} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

and may be partitioned as

$$\begin{bmatrix} J^k_{1,1} & J^k_{1,2} & ... & J^k_{1,m} \\ J^k_{2,1} & J^k_{2,2} & ... & J^k_{2,m} \\ ... & ... & ... & ... \\ J^k_{m,1} & J^k_{m,2} & ... & J^k_{m,m} \end{bmatrix} \tag{7}$$

according to all the endogenous variables q_i. Hence, each sub-matrix $J_{i,j}$, for $i, j = 1, ..., m$ is of dimension $T \times T$ with each element given by

$$\begin{aligned} J_{i,j,t,\Delta_t} \quad &= \quad \begin{cases} 1 - \frac{\partial h_{i,t}(Q^k)}{\partial q^k_{j,t+\Delta_t}} & \text{for } \Delta_t = 0 \text{ and } i = j \\ -\frac{\partial h_{i,t}(Q^k)}{\partial q^k_{j,t+\Delta_t}} & \text{else} \end{cases} \\ \text{for} \quad & t = 0, ..., T, \\ \text{where} \quad & -t \leq \Delta_t \leq T - t \end{aligned} \tag{8}$$

For a heterogeneous agent model with finite life-times of each individual agent, $-\frac{\partial h_{i,t}(Q^k)}{\partial q^k_{j,t+\Delta_t}} = 0$ for Δ_t sufficiently large. Hence $J[G(Q^k)]$ is sparse. Despite, it is generally quite costly to determine all non-zero elements of the Jacobi matrix $J[G(Q^k)]$ as T (and m) become large.

3.1 The Steady State

Suppose now that variables in Q are transformed such that they are constant in the steady state. E.g., q_1 could be a time series of the capital to output ratio and q_2 of the labor supply ratio in a closed economy growth model with endogenous labor supply ($m = 2$). Further, note that domain truncation imposes a restriction on the equation system which is mirrored by a Jacobi matrix of finite dimension and hence by the restriction on Δ_t in equation (8) requiring that $-t \leq \Delta_t \leq T-t$. This restriction is invalid if the economy is in steady state. For such a model the restriction on Δ_t is $-T_0 - t \leq \Delta_t \leq T_0 - t$, where $T_0 \leq T$, since, as noted above, $-\frac{\partial h_{i,t}(Q^k)}{\partial q^k_{j,t+\Delta_t}}$ may be zero for Δ_t sufficiently large. Further, since the elements of each $\{q_i\}_{i=1}^m$ are constant in the steady state, the partial derivatives in equation (8) will be constant across time as well. The corresponding representation of the elements of the actual Jacobi matrix in equation (8) is given by

$$J^{T_0}_{i,j,\Delta_t} = \begin{cases} 1 - \frac{\partial h_i(Q^k)}{\partial q^k_{j,\Delta_t}} & \text{for } \Delta_t = 0 \text{ and } i = j \\ -\frac{\partial h_i(Q^k)}{\partial q^k_{j,\Delta_t}} & \text{else} \end{cases} \quad \text{where } -T_0 - t \leq \Delta_t \leq T_0 - t, \tag{9}$$

which only depends on Δ_t and not on the time period t itself.

Therefore each of the m^2 different sub-matrices of the Jacobi matrix defined in equation (7) can be written as

$$J_{i,j}^{T_0} = \left(\sum_{\Delta_t=-T_0-t}^{T_0-t} D_{\Delta_t} - \frac{\partial h_i(Q^k)}{\partial q_{j,\Delta_t}^k} \right) \cdot I_{T\times T}, \text{where } \begin{cases} D_{\Delta_t} = 1 & \text{for } \Delta_t = 0 \text{ and } i = j \\ D_{\Delta_t} = 0 & \text{else} \end{cases}. \tag{10}$$

For steady state situations of the economic model the exact Jacobi matrix is accordingly given by

$$\hat{J}^k = [W^{-1}]_{(m\times m)}^k \otimes I_{(T\times T)} = \begin{bmatrix} \omega_{1,1}^k I_{(T\times T)} & \omega_{1,2}^k I_{(T\times T)} & \dots & \omega_{1,m}^k I_{(T\times T)} \\ \omega_{2,1}^k I_{(T\times T)} & \omega_{2,2}^k I_{(T\times T)} & \dots & \omega_{2,m}^k I_{(T\times T)} \\ \dots & & & \\ \omega_{m,1}^k I_{(T\times T)} & \omega_{m,2}^k I_{(T\times T)} & \dots & \omega_{m,m}^k I_{(T\times T)} \end{bmatrix}. \tag{11}$$

This structure of the Jacobi matrix is very different from a scaled identity matrix and considerably relaxes the restrictions imposed by standard Gauss-Seidel iterations. Note that m is generally small and hence W is of low dimension.

To summarize: The Jacobi matrix of a root-finding problem of an economic model as represented in equation (6) is generally quite large. For example, for a closed economy model with endogenous capital formation and endogenous labor supply ($m = 2$) that is solved for $T = 300$ years - a standard time horizon for OLG models solved at an annual frequency -, the Jacobi matrix consists of $(mT)^2 = 360,000$ elements. However, in the steady state of the model and if the elements in Q are defined such that they are constant in the steady state, the actual Jacobi matrix reduces to the Kronecker product of the low-dimensional W^{-1}-matrix and an identity matrix. Hence, for the above example, the exact Jacobi matrix effectively consists of only $m^2 = 4$ elements. This Jacobi matrix can easily be determined by standard finite difference methods in the first tatonnement iteration and can be updated by Broyden's method as the iteration proceeds, see Section 3.3 below. Hence, the final Jacobi matrix derived in steady state iterations, $J^{*,ss}$ is asymptotically optimal and convergence will be super-linear, compare Press et al. (1992).

3.2 The transition

Since T is generally quite large, transition calculations may take considerable time to compute. Against this background, the idea behind the implementation of GSQN for transition calculations is to use the Jacobi matrix derived during (fast) steady state calculations as an initial approximate Jacobi matrix for transition calculations and to update it by Broyden's method as the iteration proceeds, see Section 3.3 below.[10] The exact implementation of the algorithm during transition

[10]Note that applying different dampening factors for different time periods t is not reasonable since it would create artificial kinks in the time paths of Q.

calculations depends on the restrictions on the structure of the equation system imposed by (initial and final) steady states or (and) arbitrary initial conditions. Four different models can be distinguished:

- **Model 1**: The economy starts from an initial steady state and converges to a final steady state. The final steady state has been calculated.
- **Model 2**: The economy starts from an initial steady state and converges to a final steady state. The initial steady state has been calculated.
- **Model 3**: The economy starts from an initial steady state and converges to a final steady state. Both steady states have been calculated.
- **Model 4**: The economy starts from arbitrary initial conditions and converges to a final steady state. The initial conditions are known and the final steady state has been calculated.

Permanent structural changes are implicit in the definitions of all models. However, for temporary changes, the economy starts from the same steady state as it converges to and hence such a specification is nested in model 3.

In terms of equations the four different models can be written as follows. For ease of presentation it is assumed that $m = 1$. Recall that the variables in Q are transformed such that they are constant in the steady state.

- **Model 1**:

$$\begin{aligned} q_0 &= q_1 \\ q_1 &= h_1(Q) \\ q_2 &= h_2(Q) \\ &... \\ q_T &= q^{fss}, \end{aligned}$$

  where fss stands for final steady state.

- **Model 2**:

$$\begin{aligned} q_0 &= q^{iss} \\ q_1 &= h_1(Q) \\ q_2 &= h_2(Q) \\ &... \\ q_T &= q_{T-1} \end{aligned}$$

  where iss stands for initial steady state.

- **Model 3**:

$$\begin{aligned} q_0 &= q^{iss} \\ q_1 &= h_1(Q) \\ q_2 &= h_2(Q) \\ &... \\ q_T &= q^{fss} \end{aligned}$$

- **Model 4**:

$$\begin{aligned} q_0 &= \bar{q}_0 \\ q_1 &= h_1(Q) \\ q_2 &= h_2(Q) \\ &... \\ q_T &= q^{fss} \end{aligned}$$

For Models 1 to 2 it is assumed that the final (or initial) steady state is calculated during the transition solution while the initial (or final) steady state is already known from steady state calculations. The GSQN Jacobi matrix derived during final (initial) steady state calculations, J^{fss} (J^{iss}), is then used as initial Jacobi matrix and updated by Broyden's method using the information contained in $Q^k_{i,t^{ss}} = \{q_{i,t^{ss}}\}_{i=1}^m$ and $G(Q^k_{i,t^{ss}}) = \{g(q_{i,t^{ss}})\}_{i=1}^m$ where $t^{ss} = 1$ ($t^{ss} = T$), i.e., the information contained in the initial (final) steady state period (compare Section 3.3 below).[11] For Model 3 it is assumed that both steady states were calculated during steady state calculations. GSQN is then implemented by using the iteration matrix derived during these steady state calculations (either initial or final), J^{ss}, throughout all transition iterations. The procedure for Model 4 is equivalent to Model 3.

To summarize: While the Jacobi matrix determined by the suggested method is asymptotically optimal as Q^k approaches Q^{ss} for steady state calculations, it is a good approximation for transition calculations. The matrix W may therefore be interpreted either as an approximate Jacobi matrix or as an $m \times m$ matrix of multiple dampening factors that vary with the iteration number k, compare (Hughes Hallet 1984).

3.3 Implementation of Gauss-Seidel-Quasi-Newton iterations

This section summarizes the implementation steps of GSQN. It thereby makes explicit that an application of GSQN just requires to augment intuitive taton-

[11] Updating J^k by Broyden's method is not necessary, but using the additional information contained in each iteration step k is more efficient than using a constant approximate Jacobi matrix throughout.

nement iterations with standard and well-established numerical methods.

It is well-known that if $G(Q)$ is continuously differentiable over a convex set D containing the equilibrium values Q^* with $G(Q^*) = 0$, then there exists an open set C about Q^* such that equation (6) converges at least linearly from any $Q^0 \in C$. If in addition the Lipschitz condition $\|Q^k - Q^*\| \leq d|Q^{k-1} - Q^*\|$ holds for $Q^0 \in C$ and some $d > 0$, the rate of convergence becomes quadratic. However, if the starting values Q^0 are not within C, then Newton iterations such as equation (6) may not be convergent. In order to obtain an iteration scheme that converges for almost any starting value, it is therefore reasonable to augment the Newton iteration by a line search method to get

$$Q^{k+1} = Q^k - s^k \hat{J}^{-1}[Q^k]G(Q^k), \tag{12}$$

where s^k is a standard variable step-size parameter and $\hat{J}$ is the GSQN (approximate) Jacobi matrix. Recall that $\hat{J}^{-1}[Q^k] = W^{-1}_{(m\times m)}[Q^k] \otimes I_{(T\times T)}$. A fast algorithm for line searches is by backtracking, see e.g., Press et al. (1992). It relies on a quadratic approximation of the (unknown) objective function given by $g(Q^k) = \frac{1}{2}G(Q^k)'G(Q^k)$ and determines a step that minimizes this quadratic approximation. If the resulting step is not acceptable, then the algorithm iterates over a cubic approximation of the objective function until an acceptable step is found.

However, since $\hat{J}^k$ is not the exact Jacobian, it is not guaranteed that the line search algorithm will give a descent step direction. Hence, the Jacobian will be re-initialized (by finite difference methods) in case the line search algorithm does not return a suitable step (after a maximum of only three line search iterations or when reaching a minimum value for s^k). For transition calculations, both line search algorithm and even more re-initializing the Jacobian can be costly in terms of computational time. Therefore, restarts of iterations reset the Jacobi matrix to the initial Jacobi matrix if line searches fail during transition iterations.

Moreover, it will be useful to re-initialize the Jacobian if the updated Jacobian $\hat{J}^k$ fails to satisfy two conditions: (i) if $\hat{J}^k$ is ill conditioned[12] and (ii) if some of the elements of $\hat{J}^k$ do not satisfy certain criteria reflecting prior knowledge regarding their value. E.g., for the applications considered in Section 5, it is required that the diagonal elements of J^k are positive. Condition (i) is standard and condition (ii) would automatically be fixed in the next iteration step by the methods just described (it would result in a divergent process and hence the Jacobian would be re-computed in the next iteration step). Making use of prior information is therefore not necessary but may save iterations steps.

While the application of Broyden's method is well-established it is useful to more concisely summarize the GSQN algorithm as follows:

[12]For models where the Jacobian is ill-conditioned at equilibrium, J^k would not be further updated in case Q^k approaches Q^*. In case iterations are divergent, J^k would only be scaled by line search methods.

1. Chose some initial value Q^0 and a stopping criterion ϵ. For steady state calculations, Q^0 consist of time series of any - but reasonable - constant values and for transition calculations $Q^0 = Q^{*,s}$, i.e., the equilibrium values from steady state calculations (or other constant or non-constant values, e.g., obtained during previous transition calculations).

2. Initialize the Jacobian, $\hat{J}^0 = [W^{-1}]^0 \otimes I$. Use finite difference methods for steady state calculations and $\hat{J}^0 = \hat{J}^{*,s}$ for transition calculations, i.e. the last approximate Jacobi matrix of steady state iterations (or any other initial matrix such as a scaled identity matrix).

3. For iteration k, determine Q^{k+1} by

$$Q^{k+1} = Q^k - s^k \hat{J}^{-1}[Q^k]G(Q^k), \text{ for } s^k = 1$$

   and evaluate $G(Q^k)$ as well as

$$g(Q^k) = \frac{1}{2}G(Q^k)'G(Q^k)$$

   - If $g(Q^k) < g(Q^{k-1})$ continue with step 4, else start a line-search algorithm. Use a standard backtracking algorithm for line search that stops if $g(Q^k) < g(Q^{k-1})$, if $s^k = s^{\min}$ or a maximum number of line search iterations of only three is reached. A good choice for s^{min} is 0.1, see Press et al. (1992) for details.
   - If the line search algorithm is successful then continue with step 4, else re-initialize $\hat{J}^k$ by finite difference methods in steady state iterations and by setting $\hat{J}^k = \hat{J}^0$, re-evaluate $G(Q^k)$ as well as $g(Q^k)$ and continue with step 4.

4. If $\max(||G(Q^k)/Q^k||) < \epsilon$[13] then stop and report success, else if $\Delta Q^k > \eta$, where η is some small number, determine $[\hat{W}^{-1}]^{k+1}$ by Broyden's method as

$$[\hat{W}^{-1}]^{k+1} = [\hat{W}^{-1}]^k + \frac{(\Delta G_{t^{ss}}(Q^k_{t^{ss}}) - [\hat{W}^{-1}]^k \Delta Q^k_{t^{ss}})(Q^k_{t^{ss}})'}{(Q^k_{t^{ss}})' Q^{k-1}_{t^{ss}}},$$

   where $\Delta Q^k_{t^{ss}} = Q^k_{t^{ss}} - Q^{k-1}_{t^{ss}}$ and $\Delta G_{t^{ss}}(Q^k_{t^{ss}}) = G_{t^{ss}}(Q^k_{t^{ss}}) - G_{t^{ss}}(Q^{k-1}_{t^{ss}})$. Do not update if $\Delta Q^k_{t^{ss}} \leq \eta$. t^{ss} denotes the steady state period of the model with $t^{ss} = 1$ ($t^{ss} = T$) for the initial (final) steady state, compare Section 3.2 and $Q_{t^{ss}} = (q_{1,t^{ss}}, q_{2,t^{ss}}, ..., q_{m,t^{ss}})$.
   If

   - $[\hat{W}^{-1}]^{k+1}$ is ill-conditioned or
   - $[\hat{W}^{-1}]^{k+1}$ does not satisfy prior information regarding its structure

[13] Throughout the analysis, I use the relative error tolerance only.

then re-initialize $[\hat{W}^{-1}]^{k+1}$, otherwise proceed. Re-initialize $[\hat{W}^{-1}]^{k+1}$ by first-differences in steady state iterations and by resetting $[\hat{W}^{-1}]^{k+1} = [\hat{W}^{-1}]^0$ for transition iterations. Define $\hat{J}^{k+1} = [\hat{W}^{-1}]^{k+1} \otimes I_{T\times T}$ and continue with step 3.

3.4 Further considerations

For ease of presentation, suppose throughout this section that there is no growth and hence that all variables are constant in the steady state. The assumption underlying equation (4) is that disaggregate variables need not be updated as the iteration proceeds. This is restrictive and will not be the case for most applications. Often, important feedback effects exist between disaggregate and aggregate variables in each iteration loop.

To formalize such relationships, rewrite the system of equations in (4) to the modified system

$$\begin{aligned} P^{k+1} &= S^{-1}(Q^k, B^{d,k}) \\ B^{s,k+1} &= B^s(P^{k+1}, Q^k, B^{d,k}) \\ Q^{k+1} &= D(P^{k+1}, B^{s,k+1}, B^{d,k}) \\ B^{d,k+1} &= B^d(P^{k+1}, B^{s,k+1}, B^{d,k}). \end{aligned} \tag{13}$$

For ease of presentation, the fact that only subsets of the disaggregate variables B^s and B^d are important for the above mentioned circular relationships is ignored here. Due to the block Gauss-Seidel structure, disaggregate variables of the supply model, B^s, can be substituted out and the modified system can be more concisely written as

$$\begin{aligned} Q^{k+1} &= H_1(Q^k, B^{d,k}) \\ B^{d,k+1} &= H_2(Q^{k+1}, B^{d,k}). \end{aligned} \tag{14}$$

As an example for such disaggregate variables in an OLG context consider the shadow value of leisure in a model with endogenous labor supply, compare Section 4.2 below. For given aggregate wages and disaggregate shadow wages households determine how much labor to supply. In case constraints are violated, e.g. if leisure exceeds time endowment or if labor supply is positive even though shadow wage rates are positive, then shadow wages need to be updated, compare Auerbach and Kotlikoff (1987, p.31 and p.47). Hence, there is a feedback effect between aggregate and disaggregate variables.[14]

[14] As an alternative to updating shadow wages as outer loops proceed, the household model may be solved accurately - up to some tolerance bound - by a standard shooting algorithm requiring a number of inner loop iterations per household and per outer loop. Yet, this is not efficient since accuracy of inner loops will increase automatically as the number of outer loops increases.

Instead of applying Quasi-Newton methods to the entire system of equations in (14), GSQN proceeds as follows. First, computational stability increases if disaggregate variables are related to aggregate variables, e.g., shadow wages are linked to the overall wage level. Let $p \subset P$ denote aggregate net wages and $\{b_e^d\}_{e=0}^E \subset B^d$ denote disaggregate shadow wage rates. Then by

$$\left\{ r_e = \frac{b_e^d}{p} \right\}_{e=0}^E$$

the time path of shadow wages of each age-group n is related to the overall wage level p.

Define by $R = (r_1, r_2, ...)$ where $r_i = \{\{r_{i,n,t}\}_{n=0}^{N_i}\}_{t=0}^T$ (the number of disaggregate units n may again differ across i) the set of all variables that involve transformations of B^d and Q (or P) respectively. Substituting out variables P from these relationships, the above system of equations then rewrites as

$$\begin{aligned} B^{d,k} &= V^{-1}(Q^k, R^k) \\ Q^{k+1} &= H_1(Q^k, B^{d,k}) \\ B^{d,k+1} &= H_2(Q^{k+1}, B^{d,k}) \\ R^{k+1} &= V(B^{d,k+1}, Q^{k+1}), \end{aligned}$$

where V are all non-linear functions that transform Q and B^d to R. Note that updating of the transformed variables R without dampening translates into dampened updates of the original variables B^d.

Second, dampening of updated expressions for Q^k proceeds as before. However, the circular relationships between aggregate and disaggregate variables add further "noise" to the updating of the Jacobi matrix of the reduced sub-system of non-linear equations given by

$$G_1(Q^k) = Q^k - H_1(Q^k, B^{d,k}) = 0,$$

due to the presence of the variables B^d. The elements of the corresponding dampening factor matrix are given by

$$W^{-1} = \left\{ \sum_{\Delta_t=-T_0-t}^{T_0-t} D_{\Delta_t} - \frac{\partial h_{1,i}(Q^k, B^{d,k})}{\partial q_{j,\Delta_t}^k} - \sum_{l=1}^{L} \sum_{e=0}^{E_l} \frac{\partial h_{1,i}(Q^k, B^{d,k})}{\partial b_{j,\Delta_t,l,e}^{d,k}} \frac{\partial b_{j,\Delta_t,l,e}^{d,k}}{\partial q_{j,\Delta_t}^k} \right\}_{i,j=1}^m,$$

$$\text{where } \begin{cases} D_{\Delta_t} = 1 & \text{for } \Delta_t = 0 \text{ and } i = j \\ D_{\Delta_t} = 0 & \text{else} \end{cases}.$$

Here, L denotes the number of relevant disaggregate variables and E_l the dimension of disaggregate variable l.

For most applications the additional terms in the above expression will be small and will not be determined during finite difference evaluations of the Jacobi

matrix. Broyden's updating automatically takes into account these additional terms. The larger the additional terms, the more reasonable it will therefore be to start with any initial guess of a Jacobi matrix rather than to determine it by finite difference methods and to update it by Broyden's method as before. Dealing with disaggregate variables as described adds an additional channel through which GSQN combines first-order with second-order methods.

4 Economic examples

This section describes two economic examples to illustrate the GSQN algorithm. As a first example, the familiar hog-cycle model is used to highlight the restrictions implicit to first-order iterative schemes such as Gauss-Seidel. The second example is a conventional large scale open-economy OLG model. The OLG model is used for a simulation analysis regarding the relative performances of FGS and GSQN, respectively. Results of this simulation analysis are presented in Section 5.

4.1 The hog-cycle model

The hog-cycle model is used to highlight the restrictions implicit in first-order iterations such as Gauss-Seidel. To this end, the relationship between the approximate Jacobi matrix and the actual Jacobi matrix implied by the the economic model is reversed: the question asked here is what kind of restrictions must be imposed on the economic model such that the Jacobi matrix implied by the fixed dampening factor is the actual Jacobi matrix of the economic model.

4.1.1 One-good model

The familiar static one-good hog-cycle model consists of a demand and a supply relationship. Suppose that

$$\begin{aligned} p &= s^{-1}(q) \\ q &= d(p) \end{aligned}$$

describes these economic relationships. As before these equations may be more concisely written as

$$\begin{aligned} q &= d(s^{-1}(q)) \Leftrightarrow \\ g(q) &= q - d(s^{-1}(q)) = q - h(q) = 0 \end{aligned}$$

and the (1×1) Jacobi matrix is given by

$$J = 1 - \frac{\partial h(q)}{\partial q} = \frac{\partial d(p)}{\partial p}\frac{\partial s^{-1}(q)}{\partial q},$$

which - among other things - depends on q. But a constant dampening factor w restricts the Jacobi matrix to be independent of q which will be the case if the inverse supply function, $s^{-1}(q)$, and the demand function, $d(p)$, are linear.[15]

Suppose that

$$\begin{aligned} p = s^{-1}(q) &= a_0 + a_1 q \\ q = d(p) &= b_0 + b_1 p \end{aligned}$$

then

$$J \;=\; 1 - \frac{\partial h(q)}{\partial q} = 1 - \frac{\partial d(p)}{\partial p}\frac{\partial s^{-1}(q)}{\partial q} = 1 - b_1 a_1.$$

Then the restriction implied by $J = w^{-1}$ is of course only correct if the relationship between the slopes of the demand and supply curves satisfies

$$a_1 = \frac{1 - w^{-1}}{b_1}.$$

4.1.2 Two-goods model

Suppose that the above model is extended to a two good model and that (inverse) supply and demand functions are linear and of the following form

$$\begin{aligned} s_1^{-1}(q_1, q_2) = p_1 &= a_{10} + a_{11} q_1 + a_{12} q_2 \\ s_2^{-1}(q_1, q_2) = p_2 &= a_{20} + a_{21} q_1 + a_{22} q_2 \\ d_1(p_1, p_2) = q_1 &= b_{10} + b_{11} p_1 + b_{12} p_2 \\ d_2(p_1, p_2) = q_2 &= b_{20} + b_{21} p_1 + b_{22} p_2. \end{aligned}$$

The corresponding functions $h_1(q_1, q_2)$ and $h_2(q_1, q_2)$ are accordingly given by

$$\begin{aligned} h_1(q_1, q_2) &= b_{10} + b_{11}(a_{10} + a_{11} q_1 + a_{12} q_2) + b_{12}(a_{20} + a_{21} q_1 + a_{22} q_2) \\ h_2(q_1, q_2) &= b_{20} + b_{21}(a_{10} + a_{11} q_1 + a_{12} q_2) + b_{22}(a_{20} + a_{21} q_1 + a_{22} q_2) \end{aligned}$$

and the Jacobi matrix of the system of equations $Q - S^{-1}(D(Q)) = 0$, where $Q = (q_1, q_2)$ becomes

$$J = \begin{bmatrix} 1 - \frac{\partial h_1(q_1,q_2)}{\partial q_1} & -\frac{\partial h_1(q_1,q_2)}{\partial q_2} \\ -\frac{\partial h_2(q_1,q_2)}{\partial q_1} & 1 - \frac{\partial h_2(q_1,q_2)}{\partial q_2} \end{bmatrix} = \begin{bmatrix} 1 - (b_{11}a_{11} + b_{12}a_{21}) & -(b_{11}a_{12} + b_{12}a_{21}) \\ -(b_{21}a_{11} + b_{22}a_{21}) & 1 - (b_{22}a_{22} + b_{21}a_{12}) \end{bmatrix}.$$

If $a_{11} \neq 0$, $a_{22} \neq 0$, $b_{11} \neq 0$ and $b_{22} \neq 0$, then the off-diagonal elements of J will only be zero iff

$$a_{12} = a_{21} = b_{12} = b_{21} = 0.$$

[15] Linearity of both curves is only a sufficient condition. For example, J will also be independent of q if $p = s^{-1}(q) = \sqrt{a_0 + a_1 q}$ and $q = d(p) = b_0 + b_1 p^2$.

This condition implies that cross-price elasticities of demand are equal to zero and that supplier's prices for good i are independent of supply of good $j \neq i$.

If these conditions hold, equality of off-diagonal elements of J further implies restrictions on the relationship between demand and supply curves for each good - just as in the above one-good example -, but also on the relationship across the two goods, since then

$$\begin{aligned} 1 - a_{11}b_{11} &= 1 - a_{22}b_{22} = w^{-1} \Leftrightarrow \\ a_{11}b_{11} &= a_{22}b_{22}. \end{aligned}$$

A number of lessons can be learned from these simple examples. First, it is obvious that all these conditions imply strong restrictions on both technology and preferences and will likely not hold even for these very simple linear models. Second, the restrictions are less likely to hold if the size of the model increases, i.e., if additional markets are added. Furthermore, assume that an explicit representation of the demand and supply functions does not exist for a linear model as the one considered above. Newton based methods immediately converge for linear models once the Jacobi matrix of the system is known. If it needs to be evaluated, then GSQN would require $m^2 + 2$ iterations to calculate the equilibrium (one iteration to calculate the initial values Q^{k+1} for a given starting value Q^k, m^2 iterations to calculate the Jacobi matrix and one more iteration to calculate the final solution). In contrast, FGS only needs 2 iterations if the economic model meets the restrictions implicit in one-parameter dampening. Therefore, third, in the unlikely event that the restrictions imposed by FGS are (approximately) valid, FGS will of course converge faster than GSQN. This is the more unlikely the larger is the economic model.

4.2 The OLG simulation model

This section develops an OLG model which is used for comparison of the relative performance of FGS and GSQN in Section 5. The simulation model is a three-country version of a multi-country OLG model developed by Börsch-Supan et al. (2003) in the tradition of Auerbach and Kotlikoff (1987). Here, a simplified version with stylized demographic data is used. The analysis in Section 5 distinguishes between four alternative scenarios by increasing m from one to four: (i) one-country closed economy model with exogenous labor supply ($m = 1$), (ii) one-country closed economy model with endogenous labor supply ($m = 2$), (iii) two-country open economy model with endogenous labor supply ($m = 3$) and (iv) three-country open economy model with endogenous labor supply ($m = 4$). The macroeconomic simulation model is based on a stylized demographic model used to simulate transitions which is described next.

4.2.1 The demographic model

Demographic projections enter the simulation model via time-specific sizes of living cohorts in year t denoted by $N_{t,a,i}$ where a is age and i is the country index. Sex is irrelevant for the economic model. Cohorts face mortality risk: $\varsigma_{t,a,i}$ denotes the age and time specific conditional survival probability and $\pi_{t,a,i}$ the unconditional survival probability. There is no migration. The size of a living cohorts is determined recursively by $N_{t+1,a+1,i} = N_{t,a,i}\varsigma_{t,a,i}$ for $a = 1, ..., 80$. Each year the number of newborns is determined by age and time specific fertility rates. Birth is given between the ages $20 - 40$ where fertility rates are assumed to be constant.

For all three countries, a demographic transition is assumed lasting for 100 years. The assumptions regarding the demographic transitions are arbitrary and are only set to simulate heterogeneous transitions across countries. Departing from an initially constant total population, the transition starts in year 100 and is characterized by a steadily increasing life expectancy at birth from 50 to 60, 45 to 60 and 40 to 60 for countries one to three respectively.[16] In addition, a fertility transition is assumed for countries one and three. In country one, a baby boom is simulated for years 100 to 150 - an increase of the total fertility rate (TFR) from replacement level of about 2.1 to 3 - which is followed by a baby bust for years 151 to 200 being characterized by a steady decrease of TFR to 1. After the bust, the TFR again starts to increase to replacement level of about 2.1.[17] Country two is assumed to face exactly the opposite demographic transition, namely a baby bust followed by a baby boom. The timing of the transitional dynamics of fertility in country three is assumed to be identical to country one but the baby boom is stronger since total fertility rates are assumed to increase to 4.

4.2.2 The macroeconomic simulation model

General equilibrium of the overlapping generations model is constructed via the production sector where, given factor inputs (capital and labor), output and factor prices are determined. The production sector in each country consists of a representative firm that uses a CES production function which is identical across countries and given by

$$Y_{t,i} = F(\Omega_{t,i}, K_{t,i}, L_{t,i}) = \left(\alpha K_{t,i}^{-\theta} + (1-\alpha)\left(\Omega_{t,i} L_{t,i}\right)^{-\theta}\right)^{-\frac{1}{\theta}}, \tag{15}$$

where α is the factor share and $\beta = \frac{1}{1+\theta}$ is the elasticity of substitution between the two production factors. $K_{t,i}$ denotes the aggregate capital stock, $L_{t,i}$ the

[16] A different initial life-expectancy in each country is chosen such that variables across countries differ in the initial steady state.

[17] This is achieved by adjusting fertility levels each year such that population of age 1 is constant from year to year.

aggregate labor force and $\Omega_{t,i}$ is labor augmenting technological change (Harrod neutral) growing at a constant rate g.

From static profit maximization and by the assumption of perfect capital markets, the (world) interest rate is given by

$$r_t = \alpha \frac{Y_{t,i}}{K_{t,i}} - \delta, \tag{16}$$

and the wage rate in each country is

$$w_{t,i} = (1-\alpha)\frac{Y_{t,i}}{L_{t,i}} \tag{17}$$

In order to determine aggregate consumption, optimal household behavior is derived from intertemporal utility maximization. By choosing an optimal consumption and labor supply path, each generation, economically active from period t on, maximizes the sum of remaining discounted life-time utility taking interest rates and wage rates as given from equations (16) and (17). The economic life of a cohort begins at the age of 20, for which $a = 1$ below. The maximum economic age people can reach is denoted by $Z = 60$. For the exogenous labor supply mode, it is assumed that all households supply one unit of labor for a period of 40 years, $a = 1, ..., 40$, and are retired thereafter (supplying zero units of labor). The endogenous labor supply mode does not restrict retirement age. If agents decide to supply zero units of labor, then shadow wages are calculated, see below. Households face the risk of prematurely dying with positive wealth. To rule out accidental bequests it is assumed that one-period ahead perfect annuity markets exist which perfectly insure agents against the event of early death.

In a given period a representative cohort of age a born in year t maximizes the sum of discounted life-time utility. In order to insure a stationary steady state, utility is assumed to be of the familiar Cobb-Douglas form and is given by

$$U(C_{t,a,i}, 1 - l_{t,a,i}) = \frac{1}{1-\sigma}\left(C_{t,a,i}^{\phi}(1-l_{t,a,i})^{1-\phi}\right)^{1-\sigma}, \tag{18}$$

where $C_{t,a,i}$ is consumption, $l_{t,a,i}$ is leisure, σ is the coefficient of relative risk aversion and ϕ is a weight attached to leisure.[18]

A household born in time period t maximizes

$$\max_{\{C_{t+a,a,i}, 1-l_{t+a,a,i}\}_{a=1}^{Z}} U = \sum_{a=1}^{Z}\left(\frac{1}{1+\rho}\right)^{a-1}\pi_{t+a,a,i}U(C_{t+a,a,i}, 1-l_{t+a,a,i}), \tag{19}$$

subject to a dynamic budget constraint given by:

$$A_{t+a+1,a+1,i} = \frac{1}{\varsigma_{t+a,a,i}}\left(A_{t+a,a}(1+r_{t+a}) + w_{t+a,i}l_{t+a,a,i} - C_{t+a,a,i}\right) \tag{20}$$

[18] As Auerbach and Kotlikoff (1987) point out a steady state does not exist under CES utility if wages are growing, see also Altig et al. (2001) and Börsch-Supan et al. (2003).

where the term $\frac{1}{\varsigma_{t+a,a,i}}$ reflects the assumption of perfect annuity markets.[19] A second constraint requires

$$0 \leq l_{t,a,i} \leq 1. \tag{21}$$

Maximization yields the inter-temporal Euler equation of consumption,

$$\frac{C_{t+a+1,a+1,i}}{C_{t+a,a,i}} = (\beta(1 + r_{t+a+1}))^{\frac{1}{\sigma}} \left(\frac{v_{t+a+1,i}}{v_{t+a,i}} \right)^{\frac{1}{\sigma}}, \tag{22}$$

and the intra-temporal Euler equation between consumption and leisure,

$$1 - l_{t,a,i} = u_{t,a,i} C_{t,a,i}, \tag{23}$$

where

$$u_{t,a,i} = \frac{1-\phi}{\phi} \frac{1}{\tilde{w}_{t,a,i}}, \tag{24}$$

$$v_{t,a,i} = u_{t,a,i}^{(1-\sigma)(1-\phi)} \tag{25}$$

and

$$\tilde{w}_{t,a,i} = w_{t,i} + \mu_{t,a,i}. \tag{26}$$

$\mu_{t,a,i}$ denotes shadow wages.

Equilibrium is constructed via aggregating all household's assets and labor supply decisions in any time period t. As described above, aggregate variables which are stationary will be used to solve the fixed point problem of equation (3). Due to the assumption of perfect world capital markets, the aggregate world capital to output ratio is given by

$$k_t^y = \frac{\sum_{i=1}^{R} A_{t,i}}{\sum_{i=1}^{R} Y_{t,i}}, \tag{27}$$

where R denotes the number of countries considered and $A_{t,i} = \sum_{a=1}^{Z} A_{t,a,i}$. The aggregate labor force participation rate in any country i is given by

$$l_{t,i} = \frac{L_{t,i}}{N_{t,i}}, \tag{28}$$

where $L_{t,i} = \sum_{a=1}^{Z} l_{t,a,i} N_{t,a,i}$. In terms of notation of Section 3, the variables P and Q depend on the scenarios considered:

- Exogenous labor supply / closed economy:
  $P = \{r_t\}_{t=1}^{T}$ and $Q = \{k_t^y\}_{t=1}^{T}$

[19] By the assumption of perfect annuity markets, end of period assets of households prematurely dying with positive (or negative) wealth are equally shared by the surviving members of the same cohort.

- Endogenous labor supply / closed economy:
  $P = (\{r_t\}_{t=1}^T, \{w_t\}_{t=1}^T)$ and $Q = (\{k_t^y\}_{t=1}^T, \{l_t\}_{t=1}^T)$
- Endogenous labor supply / two-country open economy:
  $P = (\{r_t\}_{t=1}^T, \{w_{t,1}\}_{t=1}^T, \{w_{t,2}\}_{t=1}^T)$ and $Q = (\{k_t^y\}_{t=1}^T, \{l_{t,1}\}_{t=1}^T, \{l_{t,2}\}_{t=1}^T)$
- Endogenous labor supply / three-country open economy:
  $P = (\{r_t\}_{t=1}^T, \{w_{t,1}\}_{t=1}^T, \{w_{t,2}\}_{t=1}^T, \{w_{t,3}\}_{t=1}^T)$ and
  $Q = (\{k_t^y\}_{t=1}^T, \{l_{t,1}\}_{t=1}^T, \{l_{t,2}\}_{t=1}^T, \{l_{t,3}\}_{t=1}^T)$

In addition, shadow wage rates are updated by the methods described in Section 3.4.

The model is calibrated with stylized demographic data as described in Section 4.2.1. Calibration of structural parameters is described in Section 5.

5 Results for the OLG model

This section compares the relative preformance of FGS and GSQN for an actual application of the large-scale OLG model presented above. The analysis is grouped into two subsections. First, a steady state analysis is carried out to determine starting values of Q, $Q^{*,ss}$ and of J, $J^{*,ss}$, to be used for the transition analysis. Second, the performance of the two algorithms is compared for the demographic transition scenarios of Section 4.2.1. The transition analysis is carried out by using $Q^0 = Q^{*,ss}$ and $J^0 = J^{*,ss}$ as starting values. In terms of notation of Section 3.2, results reported below refer to model 2, i.e., I first solve for an initial steady state and then use the initial steady state values as initial conditions for the transition calculations.

The structural model parameters of the above OLG model are given by

$$\Psi = (\Omega_0, \alpha, \beta, \sigma, \rho, \delta, g, \phi).$$

In order to compare the performance of the algorithms, three different values parameters of a subset of these structural parameters, $\Psi_1 = (\alpha, \beta, \rho, \sigma)$, are combined with each other which results in $3^4 = 81$ different parameterizations of the OLG model per model simulation, see table 1.[20] These parameterizations reflect standard parameterizations chosen for OLG models in the literature. For steady state simulations, the starting value of the capital to output ratio is constant at three for the closed economy scenario with exogenous labor supply ($m = 1$). For all other models ($m > 1$), the steady state capital to output ratio resulting from previous models with $m - 1$ endogenous variables is used. The same procedure is adopted for the choice of starting values regarding the labor supply ratio: it

[20]Except for Ω_0 which is normalized in each iteration step by requiring the model to match arbitrary GDP levels of 100 for all countries.

is assumed constant at 0.5 for the closed economy model with endogenous labor supply ($m = 2$) and equilibrium labor supply shares resulting from previous computations are used for all subsequent models with $m > 2$. In addition, two alternative dampening factors $w_1 = 0.1$ and $w_2 = 0.3$ will be compared for FGS.

The convergence criterion ϵ is set to $1e-4$ for steady state and to $1e-3$ for transition calculations[21]. This is an arbitrary choice. The relative advantage of GSQN increases the lower the convergence criterion since it asymptotically converges at a super-linear rate whereas FGS converges at a linear rate. No convergence may occur under two cases: first, when Q^k is divergent or exhibits cyclical behavior and second, when $\max(|G(Q^{k^{max}})/Q^{k^{max}}|) > \epsilon$ for some maximum number of iteration steps k^{max}. To rule out the latter case, k^{max} is set to 200.[22] The convergence properties of the two algorithms are evaluated along two dimensions, number of cases without convergence as well as running time (average and median) as the time it takes for convergent runs (in seconds). Since running time per iteration step differs between the two algorithms, results for the number of iterations it takes until convergence are only reported for sake of completeness.[23]

5.1 The steady state analysis

Convergence results for the steady state analysis of the model are reported in table 2. The table is organized in four panels in increasing order of m. The first two rows of each section show results for FGS with $w = 0.1$ and $w = 0.3$ respectively. The third row shows results for GSQN. The last two columns of each row show the relative cases without convergence respectively. GSQN always converges whereas FGS may not converge for the higher value of the dampening factor ($w = 0.3$). The fourth and fifth row of each panel show the relation between running time (and number of iterations) between FGS ($w = 0.1$) and ($w = 0.3$) and GSQN for the convergent runs of FGS respectively. For example, column one shows the average running time it takes for convergent simulations of FGS divided by the average running time of those GSQN simulations for which FGS also converges.

Average convergence speed for $w = 0.1$ are about three times lower than GSQN when $m = 1$ and about 1.8 times lower when $m = 4$. This reduction in the relative performance of GSQN is due to additional computations required for GSQN to calculate the Jacobi matrix. One might regard these differences as

[21]Setting a lower convergence criterion for steady state simulations is reasonable since a higher degree of accuracy is required (steady state solutions are fixed during transition simulations), but of course not necessary.

[22]For the scenarios considered here, this is sufficiently high since all non-convergent cases reported below are due to cyclical or divergent behavior.

[23]Running time per outer loop differs between FGS and GSQN since GSQN requires additional iterations for evaluation of the Jacobi matrix and line searches, compare Section 3.3.

marginal. However, the resulting good initial estimates of J contribute to quite considerable differences in convergence speeds during the transition analysis, see below. For FGS with $w = 0.3$ the algorithm fails to converge in quite many cases (about 3.7 percent for $m = 1$ up to 27 percent when $m > 1$, but convergence speed (of the convergent runs) is higher than for FGS with $w = 0.1$. Hence, a higher value of the fixed dampening factor trades robustness for speed. If it converges, FGS with $w = 0.3$ is even faster on average than GSQN for $m = 4$. The table also shows median running times since some cases of difficulties in convergence may be driven by outliers, but results do not look much different according to this criterion.

5.2 The transition analysis

While these steady state results already show that FGS is clearly inferior, this may not seem very compelling since there are no non-convergent cases of FGS with $w = 0.1$ and absolute overall speed is high since steady state solutions are fast to compute. But of course convergence speeds slow down a lot if larger models are used during transition iterations. Hence, computational speed may become relevant after all.

A standard weighting matrix W derived in steady state simulations is e.g. given by

$$W_{4\times4} = \begin{bmatrix} 0.205 & -0.608 & -0.426 & -0.191 \\ 0.024 & 0.928 & -0.051 & -0.023 \\ 0.021 & -0.062 & 0.956 & -0.020 \\ 0.017 & -0.051 & -0.036 & 0.984 \end{bmatrix} \tag{29}$$

which is far from a scaled identity matrix as in FGS. This multiple dampening factor matrix results from a standard parameterization of the OlG model with $\alpha = 0.4$, $\beta = 1$, $\rho = 0.01$ and $\sigma = 2$. For this standard parameterization, all three algorithms converged. However, while GSQN took only 22.8 seconds, FGS took 64.353 (193.177) for $w = 0.3$ ($w = 0.1$).

More results for transition calculations are shown in table 3 where steady state solutions for $Q^0 = Q^{*,ss}$ and $\hat{J}^0 = \hat{J}^{*,ss}$ are used as initial conditions throughout. First, GSQN and FGS with $w = 0.1$ again always converge but the number of non-convergent cases of FGS with $w = 0.3$ quite significantly increases to roughly 38 percent for $m > 1$. Second, compared to FGS with $w = 0.1$, GSQN is roughly 3 to 7 times faster than FGS and this speed advantage strictly increases in the number m of endogenous variables Q. Third, the user may be lucky when using FGS with $w = 0.3$ for $m = 1$ since the algorithm might converge even faster than GSQN (if it converges). But for values of $m > 1$ GSQN is 3 to 5 times faster for those cases when FGS with $w = 0.3$ converges.

These results are striking and suggest to use GSQN with good starting values derived from steady state solutions of the simulation model or earlier transition

iterations since GSQN is so much superior and since it is so easy to implement. The most important aspect of GSQN is that these significant increases in running times relative to standard FGS are achieved at low costs since GSQN just combines traditional fixed-point iterations with simple Newton based methods. Therefore, existing sub-routines may be used for implementation.

6 Conclusions

This paper suggests to use Gauss-Seidel-Quasi-Newton (GSQN) instead of conventional fast Gauss-Seidel (FGS) iterations for solving heterogeneous agent models. Standard Quasi-Newton based methods (Broyden's method) are used to determine elements of a low-dimensional approximation of a Jacobi matrix for Gauss-Seidel iterations which considerably improves convergence both in terms of speed as well as robustness of the iterations. This approximate Jacobi matrix may also be interpreted as a matrix of multiple dampening factors (Hughes Hallet 1984). Through its use, the GSQN is a composite method of standard first-order and second-order tatonnnement methods. The particular attractiveness of the algorithm stems from the combination of low computational costs of conventional tatonnement methods with the speed of Newton based methods. It only requires augmenting these intuitive tatonnement methods with well-established and simple numerical methods.

The simulation analysis shows, that GSQN increases convergence speed by a factor of two to seven relative to FGS for transition simulations. This relative speed advantage strictly increases in the number of aggregate endogenous variables, m, required for tatonnement iterations. Therefore, GSQN enables a researcher to solve a larger simulation model within the same time frame as FGS needs for a smaller model. This allows the researcher to investigate much more interesting scenarios. Furthermore, computational speed is relevant for estimation and sensitivity analysis (see Ludwig (2004a) for an application).

The idea behind the algorithm - constructing a composite between fixed-point iterations and Quasi-Newton methods - can be applied to other economic models and solution procedures. As shown in Ludwig (2004b), the same idea can be used in fixed-point iterations to dampen coefficients that characterize polynomials used to solve rational expectations models by standard projection methods.

References

Altig, D., A. J. Auerbach, L. J. Kotlikoff, K. A. Smetters, and J. Walliser (2001). Simulating Fundamental Tax Reform in the United States. *American Economic Review 91*, 574–595.

Auerbach, A. J. and L. J. Kotlikoff (1987). *Dynamic Fiscal Policy*. Cambridge, MA: Cambridge University Press.

Börsch-Supan, A., A. Ludwig, and J. Winter (2003). Aging, Pension Reform and Capital Flows: A Multi-Country Simulation Model. MEA Discussion Paper 28-03, University of Mannheim.

Boucekkine, R. (1995). An Alternative Methodology for Solving Nonlinear Forward-Looking Models. *Journal of Economic Dynamics and Control 19 (4)*, 711–734.

Domeij, D. and M. Floden (2003). Population Aging and International Capital Flows. mimeo, Stockholm School of Economics.

Fair, R. C. and J. B. Taylor (1983). Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models. *Econometrica 51(4)*, 1169–1186.

Feroli, M. (2002). Capital Flows Among the G-7 Nations: A Demographic Perspective. mimeo, New York University.

Hagemann, L. and D. Young (1981). *Applied Iterative Methods*. Orlando: Academic Press.

Hughes Hallet, A. J. (1982). Alternative Techniques for Solving Systems of Nonlinear Equations. *Journal of Computational and Applied Mathematics 8 (1)*, 35–48.

Hughes Hallet, A. J. (1984). Multiparameter Extrapolation and Deflation Methods for Solving Equation Systems. *International Journal of Mathematics and Mathematical Sciences 7 (4)*, 793–802.

Judd, K. L. (1999). *Numerical Methods in Economics*. Cambridge, Massachusetts: The MIT Press.

Judd, K. L. (2002). The Parametric Path Method: An Alternative to Fair-Taylor and L-B-J for Solving Perfect Foresight Models. *Journal of Economic Dynamics and Control 26*, 1557–1583.

Judd, K. L., F. Kubler, and K. Schmedders (2000). Computational Methods for Dynamic Equilibria with Heterogenous Agents. mimeo, Stanford University.

Juillard, M. (1996). DYNARE: A Program for the Resolution and Simulation of Dynamic Models with Forward Variables Through the Use of a Relaxation Algorithm. CEPREMAP Working Paper, No. 9602, Paris, France.

Juillard, M., D. Laxton, P. McAdam, and H. Pioro (1998). An Algorithm Competition: First-Order Iterations versus Newton-Based Techniques. *Journal of Economic Dynamics and Control 22*, 1291–1318.

Laffargue, J. (1990). Resolution D'un Modele Macroeconomique Avec Anticipations Rationelles. *Annales d'Economic et Statistique 17*, 97–119.

Ludwig, A. (2004a). Looking Back in Overlapping Generations Models: How Well Do They Match the Data? mimeo, MEA, University of Mannheim.

Ludwig, A. (2004b). Solving Rational Expectations Models by Projection Methods: A Note on Fixed-Point Iterations. mimeo, MEA, University of Mannheim.

Ortega, J. M. and W. C. Rheinboldt (2000). *Iterative Solution of Nonlinear Equations in Several Variables*. Philadelphia: SIAM.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in C*. New York: Cambridge University Press.

Table 1: Calibration parameters

Parameter	**Value**		
capital share α	0.3	0.4	0.5
substitution elasticity β	0.8	1	1.2
coefficient of relative risk aversion σ	1	2	3
discount rate ρ	0.01	0.02	0.03
growth rate g		0.015	
depreciation rate δ		0.05	
consumption share ϕ		0.6	

Table 2: Convergence of FGS and GSQN for the steady state

	Running time		Iteration number		No convergence
	Mean	Median	Mean	Median	Fraction
	Closed economy, exogenous labor supply ($m = 1$)				
FGS($w = 0.1$)	4.16	3.97	20.31	20.00	0.00%
FGS($w = 0.3$)	2.18	1.31	10.69	6.50	**3.7%**
GSQN	1.36	1.29	5.25	5.00	0.00%
FGS($w = 0.1$)/GSQN	**3.07**	3.07	3.87	4.00	
FGS($w = 0.3$)/GSQN	1.61	1.02	2.04	1.30	
	Closed economy, endogenous labor supply ($m = 2$)				
FGS($w = 0.1$)	15.37	15.89	55.23	58.00	0.00%
FGS($w = 0.3$)	6.72	4.98	24.10	18	**27.16%**
GSQN	3.87	2.79	6.59	6.00	0.00%
FGS($w = 0.1$)/GSQN	**3.98**	5.69	8.38	9.67	
FGS($w = 0.3$)/GSQN	1.57	1.78	3.66	3.00	
	Two-country model, endogenous labor supply ($m = 3$)				
FGS($w = 0.1$)	23.81	23.42	44.49	44.00	0.00%
FGS($w = 0.3$)	11.25	8.53	21.03	16.00	**24.69%**
GSQN	9.46	6.31	3.37	3.00	0.00%
FGS($w = 0.1$)/GSQN	**2.52**	3.71	13.20	14.67	
FGS($w = 0.3$)/GSQN	1.09	1.35	6.24	5.33	
	Three-country model, endogenous labor supply ($m = 4$)				
FGS($w = 0.1$)	38.65	38.46	48.36	48.00	0.00%
FGS($w = 0.3$)	14.81	12.93	18.56	16.00	**24.69%**
GSQN	21.52	15.08	3.83	3.00	0.00%
FGS($w = 0.1$)/GSQN	**1.80**	2.55	12.64	16.00	
FGS($w = 0.3$)/GSQN	0.64	0.86	4.85	5.33	

Notes: FGS: Conventional fast Gauss-Seidel algorithm with one-parameter dampening. GSQN: Gauss-Seidel-Quasi-Newton algorithm. This table shows steady state convergence results of FGS and GSQN for four different scenarios with 81 model simulations each. The last two rows of each section show the relative performance of FGS and GSQN for convergent runs of FGS only.

Table 3: Convergence of FGS and GSQN for the transition

	Running time		Iteration number		No convergence
	Mean	Median	Mean	Median	Fraction
	Closed economy, exogenous labor supply ($m = 1$)				
FGS($w = 0.1$)	22.78	23.20	16.60	17.00	0.00%
FGS($w = 0.3$)	11.60	6.95	8.49	5.00	**3.7%**
GSQN	8.08	6.99	5.85	5.00	0.00%
FGS($w = 0.1$)/GSQN	**2.82**	3.32	2.84	3.40	
FGS($w = 0.3$)/GSQN	1.47	0.99	1.45	1.00	
	Closed economy, endogenous labor supply ($m = 2$)				
FGS($w = 0.1$)	50.89	43.16	28.31	28.00	0.00%
FGS($w = 0.3$)	36.13	21.43	17.98	10.00	**39.51%**
GSQN	11.84	9.26	5.84	5.00	0.00%
FGS($w = 0.1$)/GSQN	**4.30**	4.66	4.85	5.60	
FGS($w = 0.3$)/GSQN	2.87	2.31	3.08	2.00	
	Two-country model, endogenous labor supply ($m = 3$)				
FGS($w = 0.1$)	151.44	120.81	40.96	40.00	0.00%
FGS($w = 0.3$)	89.86	57.30	22.94	15.00	**37.04%**
GSQN	22.73	18.28	5.60	5.00	0.00%
FGS($w = 0.1$)/GSQN	**6.66**	6.61	7.31	8.00	
FGS($w = 0.3$)/GSQN	3.57	3.14	4.09	3.00	
	Three-country model, endogenous labor supply ($m = 4$)				
FGS($w = 0.1$)	242.19	191.12	42.95	42.00	0.00%
FGS($w = 0.3$)	172.37	92.11	28.32	16.00	**38.27%**
GSQN	34.45	27.57	5.86	5.00	0.00%
FGS($w = 0.1$)/GSQN	**7.03**	6.93	7.32	8.40	
FGS($w = 0.3$)/GSQN	4.56	3.34	4.83	3.20	

Notes: FGS: Conventional fast Gauss-Seidel algorithm with one-parameter dampening. GSQN: Gauss-Seidel-Quasi-Newton algorithm. This table shows transition convergence results of FGS and GSQN for four different scenarios with 81 model simulations each when results derived from steady state calculations are used as starting values. The last two rows of each section show the relative performance of FGS and GSQN for convergent runs of FGS only.

Discussion Paper Series

Mannheim Research Institute for the Economics of Aging Universität Mannheim

To order copies, please direct your request to the author of the title in question.

Nr.	Autoren	Titel	Jahr
45-04	Matthias Weiss	Skill-Biased Technological Change: Is there Hope for the Unskilled?	04
46-04	Hans-Martin von Gaudecker	Regionale Mortalitätsunterschiede in Baden-Württemberg	04
47-04	Karsten Hank	Effects of Early Life Family on Women's Late Life Labour Market Behaviour: An Analysis of the Relationship between Childbearing and Retirement in Western Germany	04
48-04	Christina Benita Wilke	Ein Simulationsmodell des Rentenversicherungssystems: Konzeption und ausgewählte Anwendungen von MEA-PENSIM	04
49-04	Axel Börsch-Supan, Alexander Ludwig und Anette Reil-Held	Hochrechnungsmethoden und Szenarien für gesetzliche und private Renteninformationen	04
50-04	Axel Börsch-Supan	From Public Pensions to Private Savings: The Current Pension Reform Process in Europe	04
51-04	Axel Börsch-Supan	Gesamtwirtschaftliche Folgen des demographischen Wandels	04
52-04	Axel Börsch-Supan	Mind the Gap: The Effectiveness of Incentives to Boost Retirement Saving in Europe	04
53-04	Joachim Winter		04
54-04	Axel Börsch-Supan	Aus der Not eine Tugend – Zukunftsperspektiven einer alternden Gesellschaft	04
55-04	Axel Börsch-Supan	Global Aging – Issues, Answers, More Questions	04
56-04	Axel Börsch-Supan	Was bedeutet der demographische Wandel für die Wirtschaft Baden-Württembergs?	04
57-04	Hendrik Jürges	Self-assessed health, reference levels, and mortality	04
58-04	Alexander Ludwig	Improving Tatonnement Methods for Solving Heterogeneous Agent Models	04