Appendix 1: the algorithm

Simulated annealing algorithms [5] have been found to be highly robust in locating global maxima. With large numbers of parameters both global grid search and normal multi-dimensional hill-climbing search (where combinations of parameters are varied) are virtually impossible to implement; however, it clearly may well be possible in future research to add elements from such procedures to speed up our algorithm. What our algorithm does is, starting from an initial set of parameters estimated by single equation methods[6], to vary each parameter in turn by plus and minus some percentage (of its initial value, this reference value being held constant). Whichever parameter's movement generates the biggest improvement in the likelihood is adopted; the operation is repeated for the newly altered parameter set. The search process begins with + or −10% variations in parameter values; once any improvement with such a step variation is exhausted, it tries 5% variations; then 2% then 1%, after which it stops. With two parameters this can be compared to searching sequentially parallel to the axis of each variable; this gives a zigzag movement up the likelihood hill, as illustrated with iso-likelihood curves in Figure 1. Even if the climb will be inefficient (in computer-time) in its zizgzag aspect, the fact that there is no practical alternative renders this irrelevant. As with all algorithms there is no guarantee it will find the global maximum; however plainly it will climb any slope it finds itself on locally. We check on this aspect in standard ways, such as allowing the initial parameter set to be varied randomly and at random points in the search process making random perturbations in the set of parameters reached. In particular at the end of the search for the global maximum, we restart the search from different, randomly chosen initial parameter values, and check that it reaches the same maximum. In the context of large macro models we have some general checks on acceptable parameter values available from much simulation and forecasting experience; large models' behaviour can sharply deteriorate when parameters move far away from such acceptable values.

In the last stage of our procedure we use the bootstrap to compute confidence intervals. (The errors used in this stage have all been purged of autocorrelation at the original FIML estimation stage by the inclusion of appropriate autocorrelation parameters in the model itself.) We do this by generating pseudo-samples from the vectors of estimated errors (with replacement) on a model consisting of the estimated parameters; these are used to produce new sets of FIML estimates. We have greatly increased the efficiency of the search algorithm over these new parameter

[5] There are some detailed differences between our procedure ('Alg') and simulated annealing ('Sim'). Alg always takes the optimum step; Sim will (subject to a probability function) take a sub-optimal solution on the basis that this may lead to a better one later. Alg searches for the global optimum by comparing its final solution with other parameter combinations, randomly chosen at the end of the solution process; thus treating this solution as the initial value, large step changes in each parameter value are tried in the search for a better solution. Sim's step size is a function of temperature (the optimised function), which therefore enters the search programme; Alg's step size is reduced when no better solution can be found with the current step size but the optimised function has no part of the search programme. So starting Sim with too small an initial value will always lead to a local solution ('quenching'); Alg applies the same percentage change to all coefficients regardless of their starting values.

[6] The calculation of the likelihood for a parameter set is done as follows, imposing full Rational Expectations in the usual certainty equivalence manner. In what follows K is the 'horizon', the number of periods from initial solution at which a terminal condition is set; N is sample size. We take a set of parameters, say set A. We are given the sample data at each period t which consists of exogenous variables at t, projections of these up to t+K using their (univariate) forcing processes, and lagged endogenous variables. Now using parameter set A we solve the model from t to t+K. This computes the t-period expectations of all endogenous variables and also implies the errors for t (u_t or v_t) . This operation is repeated for t=1 to t=N, thus yielding N error vectors from which the likelihood for set A is calculated. The grid search algorithm then tries a new paramater set B, redoing the above steps. And so on until the maximising set is found.

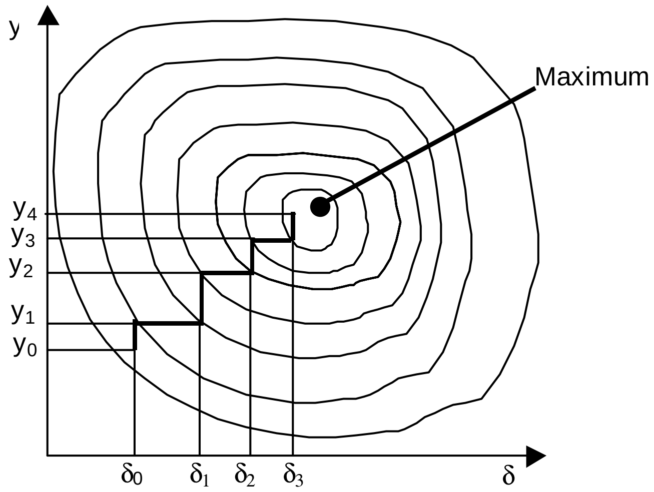

Figure 1: Hill Climbing

estimates by keeping every set of 'successful' parameters from previous bootstrap samples in the memory. For any new bootstrap data set the algorithm then checks all these sets first and chooses the best as the starting point for further search.

Appendix 2: Comparing different likelihood measures

In this appendix we present the various results when the likelihood maximand is changed, first to $\widehat{v_t}$ and then to $\widehat{y_t}$. : tables A1–4 show the ML results in each of these cases, corresponding to Tables 1 and 2 in the text. We compare the estimated coefficients for $\widehat{u_t}$, $\widehat{v_t}$ and $\widehat{y_t}$,which is done in Table A5; constants, dummies, constrained coefficients and time trends are not shown. In the last column we show the 95% confidence limits around the bootstrap mean for the $\widehat{u_t}$ estimate.We see that the differences between the first two lie outside this range only for a9 This is the unemployment term in the wage equation. It is however very small in absolute terms and the deviation is of no quantitative importance in the model.

If we compare $\widehat{y_t}$ with the other two we find again that 9 remains different. 10 is on the borderline; this is the error-correcting term in the real wage equation. The $\widehat{y_t}$ estimator is suggesting a 10% faster adjustment speed.

Plainly the estimators differ somewhat. But the differences are barely significant and indeed they are massively dwarfed by the bias in all of them. The key question about all these maximum likelihood estimators is therefore that of bias and its removal; from that viewpoint they are the same.

Table A1: Small Model Results*

	single Equation	se	FIML Est.	Bootstrp Mean	Lower 95%	Upper 95%
a2			0.38874	0.17049	-0.04092	0.42693
a3	-2.14984	0.72	-2.79479	-2.13612	-2.14984	-2.14984
a4	0.792175	0.55	0.29311	-0.14987	-0.21389	0.14259
a5	0.010303	0.0005	0.02432	0.01241	0.00567	0.01288
a6	0.803456	0.1	0.8838	0.79953	0.71508	0.80346
a19	0.271353	0.09	0.56984	0.03998	-0.21166	0.55736
a20	24.4745	0.23	24.4745	24.37035	19.5796	24.4745
a7	0.469466	0.31	0.37557	0.46447	0.32863	0.46947
a8	0.20982	0.07	0.14687	0.20803	0.16786	0.20982
a9	-0.0181	0.027	-0.01665	-0.0201	-0.02118	-0.01847
a10	-0.2238	0.06	-0.2417	-0.25318	-0.32899	-0.14995
a11	-0.28973	0.22	-0.35637	-0.26096	-0.536	0
a12	0.189295	0.12	0.28016	0.19232	-0.04732	0.45999
a21	0.1319	0.074	0.10552	0.15608	0.14509	0.16488
a13	0.164022	0.119	0	0.00466	-0.00328	0.0082
a14	0.130547	0.102	0.03916	0.02728	-0.35509	0.35509
a15	0.40656	0.0511	0	0.08152	-0.09351	0.28053
a16	-0.07581	-0.0972	0.10234	0.15661	-0.14025	0.50414
a17	0.985691	0.0107	1	1	1	1
a18	0.7		0.686	0.41845	0.063	0.644
a22	0.15		0.09	0.05398	-0.183	0.255
a23	0.019723	0.001	0.03728	0.0178	-0.00651	0.02525
a24	0.039387	0.035	0.193	0.1097	-0.06302	0.24617
a25	0.040563	0.0502	0.17929	0.17708	-0.02839	0.37521
a26	-0.01062	0.093	-0.01136	-0.05134	-0.06744	-0.03717
a27	-0.89442	0.0104	-1.0733	-0.89252	-0.89442	-0.8497
a28	0.600661		0.43248	-0.12179	-0.40845	0.35439
a29	0.001816	0.234	-0.0008	-0.00339	-0.01524	0.00948
a30	0.007839	0.231	0.00063	0.00292	-0.00902	0.01466
a31	-0.00299	0.223	-0.00299	-0.00548	-0.01779	0.00966

*From 235 bootstraps - see Annex S1 for details of model and distributions.

Table A2: Small Model, $\hat{v}_t$ in place of $\hat{u}_t$

	Single Equation	se	FIML Est.	Bootstrp	95% CI Lower	95% CI Upper
a2			0.36357	0.12561	-0.04526	0.42358
a3	-2.14984	0.72	-2.83505	-2.0965	-2.14605	-2.13493
a4	0.792175	0.55	0.34297	-0.1351	-0.21061	0.14328
a5	0.010303	0.0005	0.01779	0.004083	0.00293	0.0958
a6	0.803456	0.1	0.87232	0.80773	0.7134	0.8175
a19	0.271353	0.09	0.55565	0.06265	-0.2099	0.55267
a20	24.4745	0.23	24.52165	24.37749	19.57842	24.47743
a7	0.469466	0.31	0.32919	0.47246	0.32507	0.56473
a8	0.20982	0.07	0.15133	0.19146	0.16848	0.21121
a9	-0.0181	0.027	-0.01961	-0.02294	-0.02444	-0.01213
a10	-0.2238	0.06	-0.29396	-0.22365	-0.325	-0.14986
a11	-0.28973	0.22	-0.34119	-0.28165	-0.53942	-0.00057
a12	0.189295	0.12	0.23718	0.19757	-0.0471	0.45512
a21	0.1319	0.074	0.12629	0.14635	0.13833	0.16354
a13	0.164022	0.119	0.00461	0.03843	-0.12	0.013
a14	0.130547	0.102	0.00537	-0.00796	-0.35576	0.35039
a15	0.40656	0.0511	0.01375	0.09819	-0.09374	0.28135
a16	-0.07581	-0.0972	0.1371	0.16584	-0.14311	0.50102
a17	0.985691	0.0107	1	1	1	1
a18	0.7		0.68006	0.46458	0.06552	0.64555
a22	0.15		0.1103	0.06299	-0.18788	0.25176
a23	0.019723	0.001	0.0358	0.01875	-0.00583	0.02699
a24	0.039387	0.035	0.22846	0.15436	-0.05865	0.24588
a25	0.040563	0.0502	0.22676	0.21883	-0.02619	0.37476
a26	-0.01062	0.093	-0.01745	-0.04026	-0.06283	-0.03908
a27	-0.89442	0.0104	-1.08589	-0.90849	-0.9118	-0.85292
a28	0.600661		0.45358	-0.16935	-0.41177	0.35079
a29	0.001816	0.234	0.001061	-0.00821	-0.01391	0.01273
a30	0.007839	0.231	-0.00163	-0.00125	-0.01124	0.01367
a31	-0.00299	0.223	-0.00238	0.02445	-0.01754	0.01762

Table A3: Small Model, $\hat{v}_t$ in place of $\hat{u}_t$

	Single Equation	se	FIML Est.	Bootstrp	95% CI Lower	Upper
a2			0.36357	0.12561	-0.04526	0.42358
a3	-2.14984	0.72	-2.83505	-2.0965	-2.14605	-2.13493
a4	0.792175	0.55	0.34297	-0.1351	-0.21061	0.14328
a5	0.010303	0.0005	0.01779	0.004083	0.00293	0.0958
a6	0.803456	0.1	0.87232	0.80773	0.7134	0.8175
a19	0.271353	0.09	0.55565	0.06265	-0.2099	0.55267
a20	24.4745	0.23	24.52165	24.37749	19.57842	24.47743
a7	0.469466	0.31	0.32919	0.47246	0.32507	0.56473
a8	0.20982	0.07	0.15133	0.19146	0.16848	0.21121
a9	-0.0181	0.027	-0.01961	-0.02294	-0.02444	-0.01213
a10	-0.2238	0.06	-0.29396	-0.22365	-0.325	-0.14986
a11	-0.28973	0.22	-0.34119	-0.28165	-0.53942	-0.00057
a12	0.189295	0.12	0.23718	0.19757	-0.0471	0.45512
a21	0.1319	0.074	0.12629	0.14635	0.13833	0.16354
a13	0.164022	0.119	0.00461	0.03843	-0.12	0.013
a14	0.130547	0.102	0.00537	-0.00796	-0.35576	0.35039
a15	0.40656	0.0511	0.01375	0.09819	-0.09374	0.28135
a16	-0.07581	-0.0972	0.1371	0.16584	-0.14311	0.50102
a17	0.985691	0.0107	1	1	1	1
a18	0.7		0.68006	0.46458	0.06552	0.64555
a22	0.15		0.1103	0.06299	-0.18788	0.25176
a23	0.019723	0.001	0.0358	0.01875	-0.00583	0.02699
a24	0.039387	0.035	0.22846	0.15436	-0.05865	0.24588
a25	0.040563	0.0502	0.22676	0.21883	-0.02619	0.37476
a26	-0.01062	0.093	-0.01745	-0.04026	-0.06283	-0.03908
a27	-0.89442	0.0104	-1.08589	-0.90849	-0.9118	-0.85292
a28	0.600661		0.45358	-0.16935	-0.41177	0.35079
a29	0.001816	0.234	0.001061	-0.00821	-0.01391	0.01273
a30	0.007839	0.231	-0.00163	-0.00125	-0.01124	0.01367
a31	-0.00299	0.223	-0.00238	0.02445	-0.01754	0.01762

Table A4: **Small Model, Maximising** $L(y_t)$

	single Equation	se	FIML Est.	Bootstrp Mean	Lower 0.95	Upper 0.95
a2			0.398512	0.18157	-0.041797	0.426985
a3	-2.14984	0.72	-2.78547	-2.18562	-2.224077	-2.049028
a4	0.792175	0.55	0.35425	-0.1654	-0.223275	0.146527
a5	0.010303	0.0005	0.01458	0.01448	0.005615	0.014908
a6	0.803456	0.1	0.91254	0.8012	0.708609	0.806297
a19	0.271353	0.09	0.60154	0.042365	-0.220783	0.56517
a20	24.4745	0.23	24.4921	24.402	19.44079	24.90904
a7	0.469466	0.31	0.36254	0.44587	0.319483	0.473493
a8	0.20982	0.07	0.138654	0.198741	0.161735	0.216339
a9	-0.0181	0.027	-0.018652	-0.02254	-0.025716	-0.017692
a10	-0.2238	0.06	-0.30125	-0.2664	-0.335046	-0.143773
a11	-0.28973	0.22	-0.40225	-0.2945	-0.543076	0.007714
a12	0.189295	0.12	0.28854	0.20854	-0.048081	0.460898
a21	0.1319	0.074	0.10224	0.15423	0.139502	0.166921
a13	0.164022	0.119	0.00254	0.00432	-0.003318	0.008246
a14	0.130547	0.102	0.04125	0.029541	-0.357663	0.355576
a15	0.40656	0.0511	-0.0245	0.08812	-0.093705	0.281
a16	-0.07581	-0.0972	0.08541	0.160252	-0.143177	0.505115
a17	0.985691	0.0107	1	1	1	1
a18	0.7		0.6954	0.42012	0.062052	0.644299
a22	0.15		0.10012	0.053123	-0.185509	0.258603
a23	0.019723	0.001	0.03521	0.01802	-0.006512	0.02527
a24	0.039387	0.035	0.18452	0.162145	-0.063029	0.247111
a25	0.040563	0.0502	0.18652	0.16684	-0.028669	0.375666
a26	-0.01062	0.093	-0.021	-0.04851	-0.07851	-0.03895
a27	-0.89442	0.0104	-1.085521	-0.89314	-0.894788	-0.843911
a28	0.600661	0	0.501452	-0.1282	-0.41291	0.354563
a29	0.001816	0.234	-0.0005	-0.000389	-0.015547	0.010204
a30	0.007839	0.231	0.0007	0.00031	-0.009116	0.014738
a31	-0.00299	0.223	-0.003895	-0.0062	-0.018422	0.0104

Table A5: Comparison of coefficients

	$\hat{u}_t$	$\hat{v}_t$	$\hat{y}_t$	Upper 95%	Lower 95%
a2	0.17049	0.12561	0.18157	-0.04092	0.42693
a3	-2.13612	-2.0965	-2.18562	-2.14984	-2.14984
a4	-0.14987	-0.1351	-0.1654	-0.21389	0.14259
a6	0.79953	0.80773	0.8012	0.71508	0.80346
a19	0.03998	0.06265	0.042365	-0.21166	0.55736
a7	0.46447	0.47246	0.44587	0.32863	0.46947
a8	0.20803	0.19146	0.198741	0.16786	0.20982
a9	-0.0201	-0.02294	-0.02254	-0.02118	-0.01847
a10	-0.25318	-0.22365	-0.2664	-0.32899	-0.14995
a11	-0.26096	-0.28165	-0.2945	-0.536	0
a12	0.19232	0.19757	0.20854	-0.04732	0.45999
a14	0.02728	-0.00796	0.029541	-0.35509	0.35509
a15	0.08152	0.09819	0.08812	-0.09351	0.28053
a16	0.15661	0.16584	0.160252	-0.14025	0.50414
a18	0.41845	0.46458	0.42012	0.063	0.644
a22	0.05398	0.06299	0.053123	-0.183	0.255
a24	0.1097	0.15436	0.62145	-0.06302	0.24617
a25	0.17708	0.21883	0.16684	-0.02839	0.37521
a26	-0.05134	-0.04026	-0.04851	-0.06744	-0.03717
a27	-0.89252	-0.90849	-0.89314	-0.89442	-0.8497
a28	-0.12179	-0.16935	-0.1282	-0.40845	0.35439