Accuracy Verification for Numerical Solutions of Equilibrium Models

Leonid Kogan*   Indrajit Mitra†

March 28, 2014

Abstract

We propose a simulation-based procedure for evaluating approximation accuracy of numerical solutions of general equilibrium models with heterogeneous agents. We measure the approximation accuracy by the magnitude of the loss suffered by the agents as a result of following suboptimal policies. Our procedure allows agents to have knowledge of the future paths of the economy under suitably imposed costs of such foresight. This method is very general, straightforward to implement, and can be used in conjunction with various solution algorithms. We illustrate our method in the context of the incomplete-markets model of Krusell and Smith (1998), where we apply it to two widely used approximation techniques: cross-sectional moment truncation and history truncation.

*Sloan School of Management, MIT, and NBER, E62-636, 100 Main Street, Cambridge, MA 02142, tel: (617)253-2289, e-mail: lkogan@mit.edu.

†Sloan School of Management, MIT, E62-682, 100 Main Street, Cambridge, MA 02142, tel: (617)715-4821, e-mail: imitra@mit.edu.

1 Introduction

Cross-sectional heterogeneity among households and firms is at the heart of many important economic phenomena. Tractable aggregation of dynamic heterogeneous-agent economies usually requires restrictive assumptions, such as complete financial markets. Financial frictions, such as credit constraints, or unhedgeable sources of risk, such as idiosyncratic labor income shocks, make aggregation impossible and give rise to equilibrium models with irreducible agent heterogeneity. In such models, the cross-sectional distribution of agent characteristics (e.g., household wealth or firm capital and productivity) is a high- or infinite-dimensional state variable.

In a highly influential paper, Krusell and Smith (1998) introduced an approximation technique for handling infinite-dimensional general-equilibrium models. Their method relies on approximating the state of the economy with a low-dimensional state vector, typically keeping track of only a few moments of the cross-sectional distribution of agent-specific variables, e.g., individual wealth. This approximation method has been applied to a number of models (Heathcote, Storesletten, and Violante (2009) and Guvenen (2011) survey solution methodology and applications). A significant practical limitation of this and alternative approximation methods is that currently there is no reliable general methodology for verifying the accuracy of approximate solutions.[1] In this paper we propose a simulation-based method for computing bounds on accuracy of solutions obtained by various approximation methods.

In an approximate equilibrium, agents follow relatively simple policies that avoid the burden of solving a dynamic optimization problem with an high-dimensional state space. This implies that their policies are sub-optimal. We measure the accuracy of the approximate solution as the welfare loss suffered by the agents as a result of using such sub-optimal policies. This concept of approximation quality has the appeal of capturing the impact of agents' errors in economically meaningful terms. It also fits naturally with the interpretation of the approximate solutions suggested by Krusell and Smith. They view the approximate solution of the original model as an exact equilibrium in a near-rational economy, in which agents pursue suboptimal policies. Such suboptimal policies are plausible as a description of near-rational behavior if the welfare loss agents suffer because they fail to fully optimize is small, and thus expanding further resources on improving the policies is economically unjustifiable. This argument is in the spirit of modeling economic agents as satisficing rather than optimizing, as in Simon (1978).[2]

Our approach establishes an upper bound on the agents' welfare loss without computing the optimal policies, which is typically a prohibitively difficult task. In particular, we alter the original problem of an agent by enlarging his information set to allow for perfect knowledge of the future path of the aggregate state process of the economy while simultaneously penalizing the agent's objective for such foresight.[3] The modified problem is much more tractable than the original problem because the aggregate state of the economy in the modified problem follows a deterministic process. Moreover, if the penalty for perfect foresight is chosen properly, which we discuss in detail in the main text, the value function of the modified problem is higher, in expectation, than the value function of the original problem. We thus obtain an upper bound on the agent's welfare, while the lower bound results from following the sub-optimal policy prescribed by the approximate solution. The gap between the two bounds limits the agent's welfare loss from above. A small gap indicates that the degree of sub-optimality is economically small, and the approximate equilibrium in indeed near-rational. A large gap, in contrast, does not immediately imply that the sub-optimal policy is grossly inefficient, as it may result from the value function of the modified problem being significantly higher that the value function of the modified and the original problem.

To illustrate the potential of our method, we apply it to the original model in Krusell and Smith (1998). It is a stochastic growth model in which individual agents face uninsurable labor income risk as well as aggregate shocks productivity of capital. Krusell and Smith compute an approximate equilibrium by summarizing the cross-sectional distribution of wealth among the agents using only the average per capita level of wealth. We also approximate the equilibrium in this model using history truncation, as in Veracierto (1997) and Chien and Lustig (2010), which entails agents keeping track of a finite history of the recent aggregate productivity shocks.

We quantify the accuracy of both methods. We establish, in particular, that both approximation methods imply relatively low individual welfare loss for most initial configurations of the economy – starting in an aggregate state of relatively high likelihood with the level of the agent's capital stock within the bulk of the cross-sectional wealth distribution. Both of the approximation methods are designed to describe equilibrium dynamics when the economy is in stochastic steady state. Thus, for the calibrated model under consideration, our simulation method confirms that both approximation techniques are accurate when used for their intended purpose.

Next, we stress-test the performance of the above approximation algorithms by using them to describe the transitional dynamics in an economy perturbed away from its steady state. Starting from the steady state of the Krusell-Smith model, we consider two experiments: (i) an unanticipated five-fold permanent increase in the volatility of aggregate productivity shocks; or (ii) an unanticipated 50% reduction in capital stock of all agents in the economy. In the first case, the economy transitions to a new steady state following a permanent regime shift. In the second case, the economy reverts to the original steady state following a large transient shock.

It is a priory unclear how well the two approximation algorithms may perform under the above experiments. Both algorithms are designed to approximate the equilibrium when the economy is in steady state. The history truncation algorithm, in particular, relies on the assumption that a finite history of aggregate productivity shocks provides accurate information about the cross-sectional distribution, which does not apply to either of the two experiments. Accordingly, we find that in the first experiment the history truncation algorithm leads to welfare loss bounds that are an order of magnitude larger than those computed around steady state of the original model. However, in the second experiment, welfare loss bounds remain comparable to those around steady state.

The Krusell-Smith algorithm is applicable to both experiments in principle, since the average per capital capital stock may provide an accurate summary of the state of the economy even away from its steady state. However, there is no general guarantee of its approximation quality. Our algorithm allows us to quantify how well the Krusell-Smith algorithm performs in each case. We find that, in both experiments, welfare loss bounds increase. However they remain of modest magnitude – comparable to those obtained for the history truncation method around steady state of the original model.

Another contribution of our paper is to extend the information relaxation methodology to problems with recursive Epstein-Zin preferences. Non-separable preferences imply that the timing of the resolution of uncertainty affects agents's welfare, making the direct application of the existing information relaxation methodology infeasible – it is simply impossible to evaluate the agent's objective over a single future path of the economy. We overcome this difficulty using the variational characterization of Epstein-Zin preferences due to Geoffard (1996) and Dumas, Uppal, and Wang (2000). Because of the wide-spread use of nonseparable preference in economic models, our analysis substantially expands applicability of information relaxation methods to analysis of general equilibrium models.

The rest of the paper is organized as follows. In Section 2 we formulate the relaxed problem and outline the construction of penalty functions. To illustrate our approach, we apply it to a model for which the optimal policy is known in closed form. In Section 4 we extend our method to non-separable preferences. In Section 3, we apply our method to the Krusell-Smith model. Section 5 concludes.

2 Information relaxation

2.1 The main idea

In our analysis of approximate equilibria, we apply the information relaxation method proposed in Brown, Smith, and Sung (2010). We introduce the main ideas of this method in this section, and refer the readers to Brown et al. (2010) for full technical details.

Consider a standard finite-horizon consumption-savings problem. Time is discrete, $t = \{0, \cdots, T\}$. Each period the agent receives a random labor income which takes two possible values $\{y_H, y_L\}$. The probability of receiving y_H is p each period. The agent chooses consumption c_t and stores the rest in a risk-free asset with constant total return R. We denote the non-anticipative consumption policy by $C = (C_0(\cdot), C_1(\cdot), ..., C_T(\cdot))$.

At each date t, the agent observes the history of income shocks realized up to and including this date, denoted by $y^t = (y_0, y_1, ..., y_t)$, but not the future shocks. All feasible consumption policies must be adapted to the information structure of the agent, i.e., consumption choices are functions of the observed past histories of income shocks. Thus, making the agent's information structure explicit, $C = (C_0(y^0), C_1(y^1), ..., C_T(y^T))$.

The agent has a time-separable constant relative risk aversion utility function with curvature γ. Let w_t denote agent's wealth at the beginning of period t. The agent solves the dynamic optimization problem

$$\max_{\{C: C_t(y^t) \leq w_t\}} \mathrm{E}_0 \left[\sum_{t=0}^{T} \beta^t \frac{C_t^{1-\gamma}}{1-\gamma} \right], \tag{1}$$

where agent's wealth and consumption satisfy the dynamic budget constraint

$$w_t = (w_{t-1} - C_{t-1})R + y_t. \tag{2}$$

We denote the value function of the above problem by $V_t(w_t)$,

$$V_t(w_t) = \mathrm{E}_t \left[\sum_{t=0}^{T} \beta^t \frac{(C_t^\star)^{1-\gamma}}{1-\gamma} \right], \tag{3}$$

where $C^\star$ is the optimal consumption policy. We denote the time-0 expected utility as V_0, suppressing the explicit dependence on initial wealth.

We formulate a relaxed problem by allowing the agent to have access to information about the future realizations of income shocks. The name "information relaxation" reflects the notion that this formulation relaxes information constraints placed on the agent. Specifically, consider a *complete information relaxation*, whereby we allow the agent to condition his consumption choices on the knowledge of the entire future sequence of income shocks. To distinguish the feasible policies of the relaxed problem from those of the original problem, we denote the former by $C^{\mathcal{R}} = (C_0^{\mathcal{R}}(y^T), C_1^{\mathcal{R}}(y^T), ..., C_T^{\mathcal{R}}(y^T))$.

While providing the agent with such informational advantage over the original formulation, we impose a penalty on the objective function, designed to offset the effect of information relaxation. The penalty is a stochastic process λ_t, which depends on the consumption choices and the entire path of income shocks, y^T: $\lambda_t(c^T, y^T)$. The only requirement we impose on the penalty process is that if the consumption policy is non-anticipative, i.e., it depends only on the information available to the agent, the resulting penalty is non-positive in expectation, i.e.,

$$\mathrm{E}_0\left[\lambda_t\left(C, y^T\right)\right] \leq 0. \tag{4}$$

It is easy to see that the value function of the relaxed problem,

$$V_0^{\mathcal{R}} = \max_{\{C^{\mathcal{R}}: C_t^{\mathcal{R}}(y^T) \leq w_t^{\mathcal{R}}\}} \mathrm{E}\left[\sum_{t=0}^{T} \beta^t \frac{\left(C_t^{\mathcal{R}}\right)^{1-\gamma}}{1-\gamma} - \lambda_t(C^{\mathcal{R}}, y^T)\right], \tag{5}$$

subject to the the dynamic budget constraint

$$w_t^{\mathcal{R}} = (w_{t-1}^{\mathcal{R}} - C_{t-1}^{\mathcal{R}})R + y_t, \tag{6}$$

is at least as high as the value function of the original problem. The reason is that the consumption policy $C^{\star}$, optimal under the agent's original problem (1-2), is also a feasible policy for the relaxed problem (5-6), and the expected penalty under such policy adds a non-negative term to the agent's expected utility, according to (27). Thus, we establish that

$$V_0^{\mathcal{R}} \geq V_0. \tag{7}$$

Next, consider a feasible but suboptimal consumption policy $\widehat{C}$. Under this suboptimal policy, the expected utility of the agent is given by

$$\widehat{V}_0 = \mathrm{E}_0\left[\sum_{t=0}^{T} \beta^t \frac{(\widehat{C}_t)^{1-\gamma}}{1-\gamma}\right], \tag{8}$$

which results in a welfare loss of $V_0 - \widehat{V}_0$. To estimate the welfare loss resulting from a suboptimal strategy, we use the inequality (7) to conclude that the agent's welfare loss is bounded above by the difference between the value function of the relaxed problem (5-6) and the expected utility under the suboptimal policy $\widehat{C}_t(y^t)$:

$$V_0 - \widehat{V}_0 \leq V_0^{\mathcal{R}} - \widehat{V}_0. \tag{9}$$

We thus have a framework for computing bounds on welfare loss resulting from suboptimal strategies: define a valid penalty process for the relaxed problem, and then compare the expected utility of the agent under information relaxation (this problem is deterministic and hence much easier to solve than the original stochastic dynamic optimization problem) with his expected utility under the suboptimal policy of interest. While this formulation is rather general, it is only useful as long as the resulting bound is relatively tight, i.e., as long as the value function of the relaxed problem $V_0^{\mathcal{R}}$ is not much higher than the expected utility of the agent under the optimal consumption policy, V_0. Brown et al. show that it is possible to make the difference $V_0^{\mathcal{R}} - V_0$ arbitrarily small. In particular, they show (using backwards induction) that under an *ideal* penalty, $V_0^{\mathcal{R}} - V_0 = 0$.

Brown et al. define an *ideal* penalty as follows. The penalty is defined for each possible sequence of income shocks, and each possible sequence of consumption choices, without requiring the consumption policy to be non-anticipating. Specifically, consider an arbitrary path of income shocks y^T, and a budget-feasible positive sequence of consumption choices $c^T = (c_0, c_1, ..., c_T)$. Note that c^T is not a consumption policy, it denotes a sequence of positive real numbers representing a particular path of consumption. The corresponding values of agent's wealth $(w_0, w_1, ..., w_T)$ satisfy the dynamic budget constraint (2). Given w_t, each term in the ideal penalty, $\lambda_t^\star(c^T, y^T)$, is defined based on the value function of the original problem:

$$\lambda_t^\star(c^T, y^T) = V_{t+1}\left(\left(w_t - c_t\right)R + y_{t+1}\right) - \mathrm{E}\left[V_{t+1}\left(\left(w_t - c_t\right)R + \widetilde{y}_{t+1}\right)|c^t, y^t\right] \tag{10}$$

$y_0, ..., y_{t+1}$ in the above expression are non-random, these are the values of income shocks from the particular path of shocks y^T for which we are defining the ideal penalty, while $\widetilde{y}_{t+1}$ is a random variable, the labor income shock at time $t+1$. The second term on the right in (10) is an expectation of V_{t+1} over the possible values of $\widetilde{y}_{t+1}$, taking the realizations of income shocks $y_0, ..., y_t$, and consumption choices $c_0, ..., c_t$ as given. Thus, the second term depends only on c^t and y^t, and so the penalty $\lambda_t^\star$ depends on c^t and y^{t+1}. In particular, $\lambda_t^\star(c^T, y^T)$ depends on the contemporaneous consumption choice c_t and the future income shock y_{t+1} explicitly, and on the earlier consumption choices c^{t-1} and income shocks y^t

implicitly, through w^t and the dynamic budget constraint. Going forward, we use more concise notation for the penalty,

$$\lambda_t^\star(c^T, y^T) = V_{t+1}(w_{t+1}) - \mathrm{E}_t\left[V_{t+1}(w_{t+1})\right]. \tag{11}$$

To develop some intuition for how the penalty affects the solution of the relaxed problem, consider the dependence of the ideal penalty on the contemporaneous consumption choice c_t, as shown in Figure 1.[4] Consider the time-1 penalty $\lambda_1^\star$, which depends on c_1, y_2, and w_1, where w_1 captures the dependence of $\lambda_1^\star$ on prior consumption choices and income shocks. We consider two values of w_1, and for both plot the penalty as a function of c_1 and y_2.

In both panels, the solid line plots the penalty as a function of the consumption choice c_1 for $y_2 = y_H$. The dash-dot line plots the penalty for $y_2 = y_L$. Consider a relaxed problem, with the agent observing the time-2 income shock in advance and using this information in his time-1 consumption decision. Without the penalty, the agent can take full advantage of his knowledge of the future. In particular, if the agent knows that the time-2 income shock is high, $y_2 = y_H$, it is optimal to choose higher time-1 consumption than if $y_2 = y_L$. The ideal penalty discourages such behavior. As long as the consumption choice is non-anticipating, i.e., c_1 does not depend on y_2, the expected value of the penalty is zero (shown by the dash line), and welfare of the agent is not impaired by the penalty. However, if the agent chooses higher consumption in the $y_2 = y_H$ state relative to the $y_2 = y_L$ state, the expected penalty is positive and lowers agent's welfare. The ideal penalty is chosen so that the benefit of perfect foresight is offset by the negative effect of the penalty, and the agent finds it optimal to chose a non-anticipative consumption policy while knowing future realizations of income shocks.

Comparing the left and the right panel, we observe the effect of the prior consumption choice on the time-1 penalty. for both realizations of y_2, the penalty $\lambda_1^\star$ is larger in absolute value for $w_1 = 5$ than for $w_1 = 4$. This illustrates that the penalty is designed to discourage the agent from conditioning time-0 consumption choices on y_2. Selecting higher c_0 in the $y_2 = y_H$ state relative to the $y_2 = y_L$ state raises the expected penalty term $\lambda_1^\star$, making it positive even if the consumption choice at time 1 is non-anticipative. This illustrates the inter-temporal connections between various penalty terms and consumption choices.

An ideal penalty is as difficult to compute as the solution of the original problem. We therefore define the penalty based on an approximation to the value function:

$$\lambda_t(c^T, y^T) = \widehat{V}_{t+1}(w_{t+1}) - \mathrm{E}_t\left[\widehat{V}_{t+1}(w_{t+1})\right]. \tag{12}$$

Figure 2 shows an estimate of the upper bound on welfare loss of an agent using sub-optimal policies. The agent uses a consumption policy based on the optimal solution of the model with the probability of the high state equal to $\widehat{p}$, whereas the true probability is $p = 0.9$. We define $\widehat{V}_t$ to be the value function resulting from the agent's consumption policy. The dashed line in Figure 2 shows the upper bound on welfare loss of the agent computed using the information relaxation approach, while the solid line is the actual welfare loss. In panels A and B, $\widehat{p} = 0.895$. In panel A, we use zero penalty. With perfect foresight, the agent is able to achieve much higher welfare than under the suboptimal policy, resulting in a finite but uninformative upper bound. In panel B, we implement the penalty based on (12). The exact welfare loss relative to the optimal policy is less than 2.5% in certainty equivalent terms. Information relaxation bounds maximum welfare loss at less than 4%. Thus, panels A and B illustrate that a prudent choice of a penalty is necessary to make the upper bound informative. We repeat the same analysis for a higher quality sub-optimal policy corresponding to $\widehat{p} = 0.899$. Panels C and D show the results. The actual welfare loss in panel D is now lower, at most 0.5%, and the upper bound is tighter as well, limiting the maximum loss to less than 2%. This is to be expected, as the value function based on $\widehat{p} = 0.899$ is closer to the true value function based on $\widehat{p} = 0.895$, resulting in a better choice of a penalty.

The general information relaxation approach follows the same logic as the basic example above, with a multivariate state vector replacing the wealth of the agent as an argument in the value function, and allowing for multiple choice variables. In addition, the general approach allows for partial information relaxations, where the agent receives some but not complete information about the future. Formally, we describe the structure of the agent's information as filtration $\mathbb{F} = \{\mathcal{F}_0, \mathcal{F}_1, ...\mathcal{F}_T\}$, and the information set of the relaxed problem as a finer filtration $\mathbb{G} = \{\mathcal{G}_0, \mathcal{G}_1, ...\mathcal{G}_T\}$, where $\mathcal{F}_t \subseteq \mathcal{G}_t \subseteq \mathcal{F}_T$. Then, we define the relaxed problem under the information structure $\mathbb{G}$, and we define the penalty process as

$$\lambda_t = \mathrm{E}_t\big[V(x_{t+1})|\mathcal{G}_t\big] - \mathrm{E}_t\big[V(x_{t+1})|\mathcal{F}_t\big]\,, \tag{13}$$

where the two expectation operators above are conditional on the corresponding information sets, and x_{t+1} denotes the time-$(t+1)$ state vector (to avoid introducing more notation, we suppress the dependence of the penalty process on the choice variables and the exogenous shocks).

3 Application: imperfect insurance with aggregate uncertainty

We demonstrate the potential of the information relaxation methodology by computing bounds on welfare loss in the incomplete market model of Krusell and Smith (1998). This model is a canonical example of a model with an infinite dimensional state space. We review the model and equilibrium concept briefly and refer the reader to the original paper Krusell and Smith (1998) for details. In this section, we compute bounds on welfare loss under two alternative solution approaches. The first one, which we call moment truncation, is due to Krusell and Smith (1998) and is based on using a small number of moments as a reduced summary of the cross-sectional distribution of agents. The second approach is based on truncating the history of aggregate shocks, and was applied to different problems by Veracierto (1997) and Chien and Lustig (2010).

3.1 The model

This is an Aiyagari model (Huggett (1993), Aiyagari (1994)) with aggregate uncertainty. Time is discrete, $t \in \{0, 1, \cdots, \infty\}$. There is a continuum of agents of unit measure with identical constant relative risk aversion preferences:

$$\mathrm{E}_0\left[\sum_{t=0}^{\infty} \beta^t \frac{{c_t}^{1-\gamma}}{1-\gamma}\right]. \tag{14}$$

There is a single consumption good produced using a Cobb-Douglas production function

$$y_t = z_t k^{\alpha} l^{1-\alpha}\,, \qquad \alpha \in [0,1]\,, \tag{15}$$

where k and l are capital and labor inputs. Each period's output is partly used for consumption and partly added to the next-period capital stock, resulting in the capital accumulation constraint

$$k_t = (1-\delta)k_{t-1} + y_{t-1} - c_{t-1}\,. \tag{16}$$

All agents are exposed to the common aggregate productivity shock z_t, which takes one of two values $z = \{z_h, z_l\}$, $z_h > z_l$. Productivity shocks follow a Markov chain.

Households collect capital rent and labor income each period. Individual labor income is exposed to idiosyncratic employment shocks, ε_t. We assume that each agent supplies $\bar{l}$ units of labor if employed ($\varepsilon_t = 1$), and zero units if unemployed ($\varepsilon_t = 0$). Employment shocks are cross-sectionally independent, conditionally on the aggregate productivity shock. Thus, based on the law of large numbers, the unemployed fraction of the population depends only on the aggregate state. We denote the equilibrium unemployment rate conditional on z_h and z_l by u_h and u_l. Then the aggregate labor supply in the two states is given by $L_h = (1-u_h)\bar{l}$ and $L_l = (1-u_l)\bar{l}$ respectively.

We look for a competitive recursive equilibrium. Let $\psi_t(k,\epsilon)$ denote the cross-sectional distribution function at time t, defined over the individual capital stock and employment status. Aggregate output depends on the aggregate capital stock, $K_t = \int \psi(k,\epsilon) dk d\epsilon$, and the aggregate supply of labor. Input prices in competitive equilibrium are determined by their marginal product, hence capital rent and wage rate are given by

$$r(K_t, L_t, z_t) = \alpha z_t \left(\frac{K_t}{L_t}\right)^{\alpha-1}, \qquad w(K_t, L_t, z_t) = (1-\alpha) z_t \left(\frac{K_t}{L_t}\right)^{\alpha}. \tag{17}$$

Individuals optimize their consumption-investment policies under rational expectations about market prices, i.e., we assume that they correctly forecast the law of motion of the equilibrium cross-sectional distribution of agents, denoted by

$$\psi_t = H(\psi_{t-1}, z_{t-1})\,. \tag{18}$$

Thus, optimal individual policies depend on the cross-sectional distribution of capital.

The value function of the agents satisfies the Bellman equation:

$$\begin{aligned} V_t(k_t, \varepsilon_t, z_t, \psi_t) &= \max_{c_t \geq 0, k_{t+1} \geq 0} \left[\frac{c_t^{1-\gamma}}{1-\gamma} + \beta \mathrm{E}_t \left[V_{t+1}(k_{t+1}, \varepsilon_{t+1}, z_{t+1}, \psi_{t+1}) \right] \right] \\ \text{where} \quad k_{t+1} &= (1 - \delta + r_t)k_t + w_t \bar{l} \varepsilon_t - c_t, \\ \psi' &= H(\psi, z), \end{aligned} \tag{19}$$

The main difficulty in computing the competitive equilibrium arises because of the dependence of equilibrium prices on the cross-sectional distribution of agents. Thus, to solve for equilibrium, we must determine the law of motion in (18). To make the problem tractable, it is common to use a low-dimensional approximation to the infinite-dimensional cross-sectional distribution ψ_t. Below we outline the two leading approximation strategies.

Moment truncation

This method, introduced in Krusell and Smith (1998), approximately captures all relevant information about the cross-sectional distribution of capital by its first K moments, $\{m_1, m_2, \cdots, m_K\}$. In particular, in their analysis Krusell and Smith restrict their attention to the cross-sectional mean m_1 (we omit the sub-script and denote the distribution mean simply by m). To speed up computation further, Krusell and Smith posit an approximate log-linear law of motion for m:

$$\widehat{H} : \log m' = a^z + b^z \log m\,, \qquad z = \{z_g, z_b\}\,. \tag{20}$$

In an approximate equilibrium, the dynamics of the cross-sectional distribution of capital in the economy must conform closely to the assumed law of motion (20), which results in a low-dimensional fixed-point problem.[5]

History truncation

Another popular approximation strategy makes the problem tractable by truncating the history of aggregate shocks. The idea behind this approximation scheme is that the equilibrium distribution of capital at any time t is a function of the initial distribution and the realized sequence of aggregate shocks $\{z_0, \cdots, z_t\}$. The key approximating assumption is that the effect of the initial distribution on equilibrium is transient, and the state of the economy depends essentially on a small number (p) of recent aggregate shocks. In our analysis, we set p to three.[6]

Under both approximation schemes, agents form their forecast for future prices based on a low dimensional proxy to the cross-sectional distribution, while the actual distribution of future prices depends on the true high-dimensional cross-sectional distribution. Thus, individual policies are sub-optimal and result in welfare loss for the agents.

3.2 The relaxed problem

We compute an upper bound on individual welfare loss by relaxing the information set of the agents, as we describe in Section 2. In particular, we start with an initial cross-sectional distribution of capital across agents and draw a sequence of aggregate shocks $z_0, \cdots, z_{T-1}$ and a panel of idiosyncratic employment shocks. Next, we aggregate the approximate equilibrium policies of the agents and compute the prices – capital returns r_t and wages w_t – corresponding to the realized sequence of aggregate shocks. To minimize the gap between the value function of the relaxed problem and the value function of the original problem of the agent, we apply a partial relaxation, revealing only future aggregate shocks but not the agent's idiosyncratic employment shocks.

We define the penalty according to (13), thus

$$\lambda_t = \mathrm{E}_t\big[\widehat{V}_{t+1}(k_{t+1}, \varepsilon_{t+1}, z_{t+1}, \widehat{\psi}_{t+1})|\mathcal{G}_t\big] - \mathrm{E}_t\big[\widehat{V}_{t+1}(k_{t+1}, \varepsilon_{t+1}, z'_t, \widehat{\psi}_{t+1})|\mathcal{F}_t\big] \,, \tag{21}$$

where $\mathcal{G}_t$ denotes the information set of the agent and $\widehat{\psi}$ is the low-dimensional proxy for ψ used to obtain the approximate equilibrium solution. For the approximate equilibrium derived using moment truncation, $\widehat{\psi}$ represents the cross-sectional mean of capital holdings by agent; for the solution using history truncation, $\widehat{\psi}$ represents the truncated recent history of aggregate shocks $\{z_t, z_{t-1}, \cdots, z_{t-3}\}$. In (21), $\mathcal{G}_t$ contains z_{t+1}, but not ε_{t+1} and therefore we average over possible employed and un-employed future states. Knowledge of the transition probabilities for z and ε allows us to compute both the terms in (21) above explicitly as a function of the decision variable of the relaxed problem, $(k_{t+1}^{\mathcal{R}}, c_t^{\mathcal{R}})$. Finally, we use the budget constraint to eliminate k_{t+1}. Along a particular path, the penalty λ_t is then a function of consumption $c_t^{\mathcal{R}}$ only.

3.3 Results

We carry out our baseline analysis using the same parameters as in Krusell and Smith (1998). The preference parameters are $\beta = 0.99$, and $\gamma = 1$. On the production side, the parameters are: the capital share $\alpha = 0.36$, the depreciation rate $\delta = 0.025$, aggregate productivity shocks take values $z_h = 1.01$, $z_l = 0.99$, and the corresponding aggregate unemployment rates are $u_h = 0.04$, $u_l = 0.10$. The transition probability matrix for the Markov chain describing the dynamics of aggregate shocks is included in the Appendix.[7]

All of our simulation results use a sample cross-section of $N = 10^4$ agents. Sample paths are $T = 10^3$ long, and we average over 100 paths to compute unbiased estimates of the value functions $\widehat{V}$ and $V^{\mathcal{R}}$. Throughout, we report the welfare loss as a fractional certainty equivalent η

$$\eta = \frac{k_0' - k_0}{k_0'}, \quad \text{where} \quad k_0' = \widehat{V}_0^{-1}\left(V_0^{\mathcal{R}}(k_0)\right). \tag{22}$$

In the equation above, $\widehat{V}$ is the expected utility under the sub-optimal policy of the agent in the approximate equilibrium, and $V^{\mathcal{R}}$ is the value function of the relaxed problem. The numerator of η is the extra capital needed by an agent following a sub-optimal policy to obtain the level of expected utility equal to the value function of the relaxed problem with initial capital $k_0^{\mathcal{R}}$, keeping all other state variables the same.

Baseline results

The welfare loss of an agent depends on the current state – the agent's capital stock, employment status, and the state of the aggregate economy – the distribution of capital across the agents. Our baseline results report welfare loss for an agent in a typical state of the economy, drawn from the stochastic steady state. Specifically, we initiate the model with a cross-sectional distribution $\psi_{-\infty}(k) \sim \mathcal{N}(k_{ss}, 3)$, where k_{ss} is the steady-state level of capital in the absence of aggregate shocks and with $L = 1$ (our results are not sensitive to the choice of the initial distribution $\psi_{-\infty}(k)$). We then draw 100 independent random paths of aggregate and individual shocks. We choose each path to be $1,000$ periods long to allow the economy to reach its stochastic steady state. We then estimate an equal-weighted average of welfare losses across the simulated paths, which represent the expected welfare loss for an agent conditional on starting in stochastic steady state. Figure 3 summarizes the results. Panel A, which uses the penalty defined by (21), shows that individual welfare losses are small, especially for high levels of initial capital. The zero penalty formulation used to produce panel B results in an uninformative upper bound. Again, the comparison of the two panels illustrates the importance of the formulation of the penalty process.

Figure 4 shows the results for the approximate solution based on history truncation. Again we find that the approximate solution implies small welfare losses for most agents in the economy. Thus, our results verify that both the original approximation of Krusell and Smith based on moment truncation, and the alternative solution using history truncation, produce approximate equilibria in which agents come very close to fully optimizing their objectives when the economy is in a typical initial state.

Transitional dynamics

Next, we consider how accurately the approximate solutions describe the transitional dynamics of the economy following an unanticipated aggregate shock. Transitional dynamics of equilibrium models is often of great interest, yet the standard solution methods are not intended to approximate equilibrium dynamics accurately when the economy is away from its steady state. It is therefore unclear a priory how well the two approximation algorithms may perform under the above experiments, and it is essential to quantify approximation accuracy when drawing conclusions about transitional dynamics of the economy based on numerical solutions.

We compute transitional dynamics of the economy following two kinds of unanticipated shocks. The first shock is a permanent regime change: the economy gradually transitions from its baseline equilibrium to the new stochastic steady state following a five-fold increase in the volatility of aggregate shocks z. We compute agents policies under the new parameter values using both moment truncation and history truncation. Under the moment truncation approximation, we assume that agents immediately switch to new policies following a regime change. For history truncation, agents switch after three periods. In our second experiment, the economy experiences a large transitory shock: a sudden loss of 50% of capital stock of every agent.

We first use the information-relaxation algorithm to quantify how well the Krusell-Smith algorithm performs in each case. Figure 5 and Figure 6 show the results of the two experiments. We find that in both experiments the bounds on individual welfare loss increase from the baseline case. The increase is particularly significant in the second case, with the bound on certainty equivalent loss increasing by an order of magnitude. The absolute magnitude of welfare loss under transitional dynamics is modest in the first experiment and is larger in the second experiment. Thus, our information relaxation method establishes that, following an increase in productivity variance, the moment truncation approximation generates relatively low individual welfare loss. for the transitional dynamics following a large loss in capital, the welfare loss due to using Krusell-Smith policies is potentially larger.

History truncation relies on the assumption that the past few aggregate productivity shocks provide accurate information about the cross-sectional distribution in the economy. Therefore, history truncation may perform poorly under transitional dynamics. We summarize the results in Figure 7 and Figure 9. In line with our intuition, history truncation algorithm leads to very large welfare loss bounds for the case of permanent shock to output variance. As we discuss in Section 2, these bounds are one-sided, and potential welfare loss suggested by our bounds may not necessarily imply that the true welfare loss is large. In this case, we confirm that agents in the model do suffer significant welfare losses when following the suboptimal strategies based on history truncation. We consider an alternative feasible suboptimal policy, which we computed as a part of the equilibrium using the Krusell-Smith algorithm. Such a policy leads to significant welfare gains for the agents relative to the native policies prescribed by the history-truncation equilibrium. Figure 8 shows the results.

For the case of transitory shock to the capital stock, welfare loss bounds increase but remain comparable to those for the economy in stochastic steady state. Thus, while not designed for this type of problem, history truncation approximation performs relatively well.

4 Non-separable preferences

In many models, preferences of agents are non-separable across time, a popular formulation being the Epstein-Zin model of preferences. Non-separable preferences pose a challenge for the information relaxation approach in its original form. Because agents with non-separable preferences care about the timing of resolution of uncertainty, their utility depends directly on the structure of the information set, and hence it is not straightforward to define the appropriate objective function for the relaxed problem. To extend the information relaxation approach to Epstein-Zin preferences, we use the time-separable dual formulation of Geoffard (1996) and Dumas et al. (2000).

We illustrate our approach for the consumption-savings problem introduced in Section 2. We discuss the general formulation in the Appendix. Instead of the time-separable preference specification, we assume that the agent has Epstein-Zin preferences and solves the recursive maximization problem

$$V_t(w_t) = \max_{c_t \in (0, w_t)} \frac{1}{\gamma} \left[c_t^{\rho} + \beta \big(\gamma \mathrm{E}_0[V_{t+1}(w_{t+1})] \big)^{\frac{\rho}{\gamma}} \right]^{\frac{\gamma}{\rho}}, \tag{23}$$

where W is the time aggregator. The parameter ρ is related to the elasticity of inter-temporal substitution by $\mathrm{EIS} = 1/(1-\rho)$, and the parameter γ is related to the coefficient of relative risk-aversion by $\mathrm{RRA} = 1 - \gamma$. β is the time-preference parameter. The agent's consumption and savings decision is subject to the budget constraint

$$w_{t+1} = (w_t - c_t)R + y_{t+1}\,. \tag{24}$$

The dual time-separable formulation of Geoffard (1996) and Dumas et al. (2000) converts (23) to a time-separable min-max problem

$$V(w_0) = \max_{\{C: c_t \le w_t\}} \min_{\nu_t} E_0 \left[\sum_{t=0}^{T} (1-\nu_t)^t F(c_t, \nu_t) \right], \tag{25}$$

where the felicity function F is the Legendre transform of the time aggregator in (23)[8]

$$F(c,\nu) = \frac{c^\gamma}{\gamma} \left[1 - \beta^{-\frac{\gamma}{\rho-\gamma}} \left(1-\nu \right)^{\frac{\rho}{\rho-\gamma}} \right]^{\frac{\rho-\gamma}{\rho}}.$$

ν_t is a stochastic discount rate process chosen to minimize the discounted sum of current and future felicities.

To produce an upper bound on welfare $V_0^{\mathcal{R}}$, the first step is to choose any feasible discount rate process $\widehat{\nu}_t(c_t)$, which is adapted to the information structure of the agent. Once this choice has been made, the relaxed problem reduces to the one in Section 2, and we follow exactly the same steps. In the notation of Section 2, the value function of the relaxed problem is

$$V_0^{\mathcal{R}}(w_0) = \max_{\{C^{\mathcal{R}}: c_t^{\mathcal{R}} \le w_t^{\mathcal{R}}\}} \mathrm{E} \left[\sum_{t=0}^{T} (1 - \widehat{\nu}_t(c_t^{\mathcal{R}}))^t \Big(F\big(c_t^{\mathcal{R}}, \widehat{\nu}_t(c_t^{\mathcal{R}})\big) - \lambda_t(C^{\mathcal{R}}, y^T) \Big) \right], \tag{26}$$

where $C^{\mathcal{R}} = (c_0^{\mathcal{R}}(y^T), c_1^{\mathcal{R}}(y^T), ..., c_T^{\mathcal{R}}(y^T))$ denote feasible consumption policies of the relaxed problem. The penalties λ_t satisfy the same condition as in Section 2

$$\mathrm{E}_0 \left[\lambda_t \left(C, y^T \right) \right] \le 0. \tag{27}$$

We solve the deterministic problem Eq. 26 for each simulated path of income shock, and average over many paths to obtain an unbiased estimate of $V_0^{\mathcal{R}}$. In the Appendix, we prove that $V_0^{\mathcal{R}} \ge V_0$.

As an explicit example of a sub-optimal policy, we consider policies $\widehat{c}_t(w_t)$ that would be optimal for an agent assigning incorrect probabilities to income shocks, i.e. using $\widehat{p}$ instead of p. Figure 10 shows welfare loss in certainty equivalent units for $\widehat{p} = 0.895$, while Figure 11 shows welfare loss for $\widehat{p} = 0.899$. The true probability of the high state is $p = 0.9$. In both plots, the dashed curve represents an upper bound to welfare loss, while the solid curve shows the actual welfare loss. The difference between the solid and the dashed curves is the difference between the value function of the relaxed problem and the true value function of the agent. This difference is smaller for $\widehat{p}$ close to p because the penalty for perfect foresight is closer to the ideal penalty, and the discount rate $\widehat{\nu}$ is closer to being optimal.

5 Conclusion

Our analysis shows that information relaxation techniques can be effectively used to establish the accuracy of approximate solutions for equilibria in heterogeneous-agent models. This methodology is general, easy to implement, and has a wide range of potential applications beyond the scope of this paper. For instance, information relaxations could be used to evaluate the accuracy of solutions obtained using perturbation techniques. The latter approach is widely used for DSGE models (for a recent application, see Mertens and Judd (2012) and Mertens (2011)) because of its ability to handle models with high-dimensional state vectors, and is supported by the computational software Dynare.

Another promising application area is equilibrium models with heterogeneous information, or higher-order beliefs. Agents in such models face high-dimensional inference and optimization problems. Common solution techniques simplify the inference problem by truncating either the history of prices or the expectation-of-expectation loop. The information relaxation approach could be used to shed light on the accuracy of such approximate solutions. Yet another natural application is to evaluating welfare loss resulting from heuristic policies motivated by behavioral biases of the agents.

Finally, as we note in the introduction, our objective is to establish near-rationality of individual policies under approximate solutions of equilibrium models, as measured by the associated welfare loss. A small welfare loss implies that individual agents have little to gain by refining their strategies. However, there is no guarantee that price dynamics in such a near-rational economy is similar to that in an exact equilibrium of the original model. An important and challenging task for future research is to develop general quantitative tools for evaluating the effect of small deviations from individual rationality on equilibrium price dynamics.

References

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics 109*(3), 659 – 684.

Akerlof, G. A. and J. L. Yellen (1985). Can small deviations from rationality make significant differences to economic equilibria? *The American Economic Review 75*, 708 – 720.

Andersen, L. and M. Broadie (2004). Primal-dual simulation algorithm for pricing multidimensional American options. *Management Science 50*(9), 1222 – 1234.

Back, K. and S. R. Pliska (1987). The shadow price of information in continuous time decision problems. *Stochastics 22*(2), 151 – 186.

Bismut, J. M. (1973). Conjugate convex functions in optimal stochastic control. *Journal of Math. Anal. Appl. 44*, 384 – 404.

Brown, D., J. E. Smith, and P. Sung (2010). Information relaxations and duality in stochastic dynamic programs. *Operations Research 58*(4), 785 – 801.

Chien, Y. and H. Lustig (2010). The market price of aggregate risk and the wealth distribution. *Review of Financial Studies 23*(4), 1596 – 1650.

Davis, M. H. A. and I. Karatzas (1994). A deterministic approach to optimal stopping. In F. P. Kelly (Ed.), *Probability, Statistics and Optimisation*, pp. 455–466. John Wiley & Sons Ltd.

Dumas, B., R. Uppal, and T. Wang (2000). Efficient intertemporal allocations with recursive utility. *Journal of Economic Theory 93*, 240 – 259.

den Haan, W. J., K. L. Judd, and M. Juillard (2010). Computational suite of models with heterogeneous agents: Incomplete markets and aggregate uncertainty. *Journal of Economic Dynamics and Control 34*(1), 1 – 3.

den Haan, W. J. and A. Marcet (1994). Accuracy in simulations. *Review of Economic Studies 61*(1), 3 – 17.

Geoffard, P. (1996). Discounting and optimizing: Capital accumulation problems as variational minmax problems. *Journal of Economic Theory 69*, 53 – 70.

Guvenen, F. (2011). Macroeconomics with heterogeneity: A practical guide. Working Paper 17622, National Bureau of Economic Research.

Hassan, T. A. and T. M. Mertens (2011, May). The social cost of near-rational investment. NBER Working Papers 17027, National Bureau of Economic Research, Inc.

Haugh, M. B. and L. Kogan (2004). Pricing American options: A duality approach. *Operations Research 52*(2), 258 – 270.

Heathcote, J., K. Storesletten, and G. L. Violante (2009). Quantitative macroeconomics with heterogeneous households. *Annual Review of Economics 1*(1), 319–354.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control 17*(5-6), 953 – 969.

Jones, S. R. G. and J. H. Stock (1987, March). Demand disturbances and aggregate fluctuations: The implications of near rationality. *Economic Journal 97*(385), 49–64.

Krusell, P. and A. A. Smith (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy 106*(5), 867 – 896.

Mertens, T. M. (2011). Excessively volatile stock markets: Equilibrium computation and policy analysis. *Working Paper*.

Mertens, T. M. and K. L. Judd (2012). Equilibrium existence and approximation for incomplete market models with substantial heterogeneity. *Working Paper*.

Merton, R. C. (1971). Optimal consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory 3*, 373 – 413.

Naish, H. F. (1993, January). The near optimality of adaptive expectations. *Journal of Economic Behavior & Organization 20*(1), 3–22.

Pliska, S. R. (1982). Duality theory for some stochastic control models. *Lecture Notes in Control and Information Sciences 43*(2), 329 – 337.

Rockafellar, R. T. and R. J. B. Wets (1976). Nonanticipativity and l^1 martingales in stochastic optimization problems. *Mathematical Programming Studies 6*, 170 – 187.

Rogers, L. C. G. (2002). Monte Carlo valuation of American options. *Mathematical Finance 12*(3), 271 – 286.

Rogers, L. C. G. (2007). Pathwise stochastic optimal control. *SIAM J. Control Optim. 46*(3), 1116–1132.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics 51*(3), 239 – 246.

Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review 68*, 1 – 16.

Veracierto, M. (1997). Plant-level irreversible investment and equilibrium business cycles. *Discussion Paper, Federal Reserve Bank of Minneapolis 115*.

Notes

[1]Two heuristic approaches are commonly used to evaluate approximate solutions. The first approach examines the accuracy of forecasts of the future state of the economy made by the parsimonious model. Forecasts of aggregate states of the economy are compared with actual realizations from the simulation, and accuracy of the forecasts, e.g., their R^2 , are used to judge the approximation quality. The limitation of this approach is that a high forecast R^2 does not guarantee low individual welfare loss. The second approach, due to den Haan and Marcet (1994), considers Euler equation errors under the approximate solution along the simulated path of the economy. Under the null hypothesis that the agent's policies are optimal, the $\mathcal{L}^2$ norm of the Euler equation errors is distributed as a χ^2 random variable, and a standard hypothesis test can be carried out. The limitation of this method is that small Euler equation errors do not imply low welfare loss. As we show in the Appendix, Euler equation errors can be small while sub-optimal policies are infinitely costly in welfare terms.

[2]The focus on individual welfare loss leaves open an important and challenging question of how closely the equilibrium policies and prices in the near-rational equilibria correspond to those in exact equilibria, in which all agents optimize their objectives fully. It is well know that small mistakes by individual agents may potentially lead to large differences in equilibrium outcomes (e.g., Akerlof and Yellen (1985), Jones and Stock (1987), Naish (1993), Hassan and Mertens (2011)).

[3]The basic idea of using information relaxations and martingale multipliers to formulate a dual stochastic optimization problem can be traced back to Bismut (1973) (in a continuous-time setting) and Rockafellar and Wets (1976) and Pliska (1982) (in the discrete-time finite horizon setting). Back and Pliska (1987) apply this technique to basic theoretical problems in financial economics. Most of the existing applications of information relaxations deal with the optimal stopping problems, typically in the context of pricing American or Bermudan options, e.g., Davis and Karatzas (1994), Rogers (2002), Haugh and Kogan (2004), and Andersen and Broadie (2004). Rogers (2007) and Brown et al. (2010) extend the information-relaxation idea to general dynamic optimization problems. We use the formulation in Brown et al. (2010), which incorporates both perfect and partial information relaxations and derives penalty processes from the value function of the ooriginal problem. Our paper is the first to apply the information relaxation approach to approximate solutions of heterogeneous-agent equilibrium models.

[4]We parameterize the problem as follows. The risk-free interest rate $R = 1.02$. Labor income follows a two-state Markov chain with $y_L = 1.0$, and $y_H = 4.0$ with independent draws each period. The probability of y_H is $p = 0.9$. The agent has power utility with relative risk aversion of 5, and time preference parameter $\beta = 0.9$. Time horizon is $T = 100$.

[5]To solve the fixed-point problem, we start with an initial guess for the four constants $\{a^z, b^z\}$. The individual optimization problem is solved for optimal policies $k'(k, \epsilon, z, m)$. With these policies, we simulate a long time series of the cross-sectional distribution using a sample cross-section of $N = 10^4$ agents. Least squares regression estimates of $\{a^z, b^z\}$ are computed based on Eq. 20. We re-solve the individual problem with these update estimates of $\{a^z, b^z\}$, and the new optimal policies are used to simulate a new time-series of m. This is used to update $\{a^z, b^z\}$, and this process is continued till the system converges. We stop when the maximum change in the policy function k' between successive iterations is less than 10^{-8}, and the change in the norm of the vector $\{a^z, b^z\}$ between successive iterations is less than 10^{-8}.

[6]In our simulations we use $T = 10^3$ and a finite cross-section of $N = 10^4$ agents. To solve for individual policies and the transition function H, we start with an initial guess for aggregate capital $K(z_{t-3}, z_{t-2}, z_{t-1}, z_t)$,

from which we compute r_t and w_t using (17). Since the aggregate shock z takes on only two values, r_t and w_t are defined over 16 points. Individual policies additionally depend on the agent's current capital and employment status. We then solve the individual decision problem (19) by backward induction. To update the function K from one iteration to the next, we simulate a long path of the cross-sectional distribution using the individual policies and obtain the average K over each of the 16 aggregate states. This provides us with an updated estimate of the prices r_t and w_t in each of the 16 states. We then re-solve the individual problem with these updated estimates, and continue the iterative process until the norm of the difference of r between two successive iterations is less than 10^{-8}.

[7]When solving the individual Bellman equation, we replace the continuous variable $0 \leq k \leq \infty$ by a finite grid with $n_k = 40$ points between 0 and $\bar{k} = 500$. Since the policy is more non-linear near $k = 0$, for better accuracy, we place more points for lower values of k using a density which is a triple exponential function. The policy function for $k > \bar{k}$ is obtained by linear extrapolation. The upper limit $\bar{k}$ is chosen with the following consideration. In a model without aggregate uncertainty ($z_h = z_l = 1$), and $L = 1$, the steady state value of capital is $k_{\text{ss}} = \left(\frac{\frac{1}{\beta} - (1-\delta)}{\alpha}\right)^{\frac{1}{\alpha - 1}}$, which for our parameter choice is $k_{\text{ss}} = 37.99$. $\bar{k}$ is chosen to be well above this value.

[8]For Epstein-Zin preferences, the time aggregator $W(c, u) = \frac{1}{\gamma}\left[c^\rho + \beta(\gamma u)^{\rho/\gamma}\right]^{\gamma/\rho}$ from (23). The Legendre transform is defined as $F(c, \nu) = \max_u \left[W(c, u) - (1 - \nu)u\right]$.

Appendix

A Euler equations and welfare loss: an example

The following example highlights the distinction between Euler equation errors and welfare loss. Consider the consumption and portfolio choice problem of an infinitely lived agent with CRRA preferences (Samuelson (1969) and Merton (1971)). The agent has access to a single risk-free asset and a single risky asset with independent and identically distributed returns. Optimal consumption and portfolio policies solve the Bellman equation

$$V(W_0) = \max_{C_t,\phi_t} E_0\left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma}}{1-\gamma}\right], \quad W_{t+1} = (W_t - C_t)\Big(R_f + \phi_t(\tilde{R}_{t+1} - R_f)\Big),$$

where β is the time-preference parameter, γ is the coefficient of relative risk aversion, R_f and $\tilde{R}$ are gross returns of the risk-free and risky asset, respectively. The optimal portfolio policy in this setting is a constant ϕ which maximizes the certainty equivalent of next period wealth (see Samuelson (1969)):

$$B^{1-\gamma} = \max_{\phi} E_t\left[\Big(R_f + \phi(\tilde{R}_{t+1} - R_f)\Big)^{1-\gamma}\right].$$

Optimal consumption is a constant fraction of wealth $C_t = \zeta^* W_t$.

As suboptimal policies, consider budget-feasible policies with a constant consumption-wealth ratio $\zeta = C_t/W_t$. Assume $\gamma > 1$. The expected utility under such policies, indexed by ζ, is given by

$$U(W_0, \zeta) = E_0\left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma}}{1-\gamma}\right] = \frac{W_0^{1-\gamma}\zeta^{1-\gamma}}{1-\gamma} \sum_{t=0}^{\infty} \beta^t (1-\zeta)^{(1-\gamma)t} B^{(1-\gamma)t}.$$

While the optimal choice is $\zeta^* = 1 - \Big(\beta B^{1-\gamma}\Big)^{1/\gamma}$, there is a critical value $\zeta_{\text{crit}} = 1 - \frac{\beta^{\frac{1}{\gamma-1}}}{B}$, such that for any $\zeta > \zeta^*$ the expected utility of the investor is infinitely negative.

The Euler residuals remain finite even with infinite loss in utility. The Euler equation errors are

$$\varepsilon_t = 1 - E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{t+1}^i\right] = 1 - E_t\left[\beta\Big((1-\zeta)(R_f + \phi(\tilde{R}_{t+1} - R_f))\Big)^{-\gamma} R_{t+1}^i\right], \qquad i = 1, 2,$$

where $R_{t+1}^1 = R_f$, and $R_{t+1}^2 = \tilde{R}_{t+1}$ is the return on the risky asset. Since

$$E_t\left[\beta\Big((1-\zeta^*)(R_f + \phi(\tilde{R}_{t+1} - R_f))\Big)^{-\gamma} R_{t+1}^i\right] = 1, \tag{28}$$

we conclude that

$$\varepsilon_t = 1 - \left(\frac{1-\zeta}{1-\zeta^*}\right)^\gamma . \tag{29}$$

Thus, the Euler equation errors are finite for $\zeta^* < \zeta < 1$, while the utility loss is infinite.

B Information relaxation for non-separable preferences

The standard formulation uses a non-linear time aggregator W to define an agent's preference recursively: $V(x_t) = W\big(c_t, \mathrm{E}_t\big[V_{t+1}(x_{t+1})\big]\big)$. The agent maximizes the utility index over the set of feasible consumption policies, which we denote by $A^{\mathbb{F}}$ to emphasize that the agent's policies must be adapted to filtration $\mathbb{F}$. The dual formulation of Geoffard (1996) and Dumas et al. (2000) casts this recursive problem into a min-max time-separable form

$$V(x_0) = \max_{C\in A^{\mathbb{F}}} \min_{\nu_t} E_0\left[\sum_{t=0}^{T}(1-\nu_t)^t F(c_t,\nu_t)\right], \quad x_{t+1} = f_t(x_t,\phi_t,\omega)\,, \tag{30}$$

where x_t is the time t state vector, ω is an element of the underlying probability space, ϕ_t is the vector of decision variables, and f is the law of motion of the state vector. The felicity function F is the Legendre transform of the time aggregator W

$$F(c,\nu) = \max_u \left[W(c,u) - (1-\nu)u\right],$$

where $u_t = \mathrm{E}_t\big[V_{t+1}(x_{t+1})\big]$. The stochastic discount rate process ν_t is chosen to minimize the discounted sum of current and future felicities

$$\nu_t^\star(c_t) = \arg\min_{\nu_t}\left[F(c_t,\nu_t) + (1-\nu_t)u_t\right].$$

An upper bound on welfare $V_0^{\mathcal{R}}$ is obtained by setting the discount rate process to a feasible, adapted process $\widehat{\nu}_t(c_t)$, and relaxing the agent's information set from the agent's information set $\mathbb{F} = \{\mathcal{F}_0, \mathcal{F}_1, ...\mathcal{F}_T\}$ to a finer filtration $\mathbb{G} = \{\mathcal{G}_0, \mathcal{G}_1, ...\mathcal{G}_T\}$, where $\mathcal{F}_t \subseteq \mathcal{G}_t \subseteq \mathcal{F}_T$. The value function of the relaxed problem is

$$V_0^{\mathcal{R}}(x_0) = \max_{C\in A^{\mathbb{G}}} \mathrm{E}_0\left[\sum_{t=0}^{T}(1-\widehat{\nu}_t(c_t))^t\Big\{F(c_t,\widehat{\nu}_t(c_t)) - \lambda_t(c_t)\Big\}\right], \tag{31}$$

where the penalties λ_t are chosen from the feasible set

$$\mathcal{L}_{\mathbb{F}} = \{\lambda \in \mathcal{L} : \mathrm{E}_0\big[\lambda(C,\omega)\big] \le 0\,, \quad \forall C \in A^{\mathbb{F}}\}\,, \tag{32}$$

Theorem 1 *The value function of the relaxed problem is an upper bound on the value* *function of the original problem,*

$$V_0(x_0) \leq V_0^{\mathcal{R}}(x_0)\,,$$

where

$$V_0^{\mathcal{R}}(x_0) = \max_{C \in A^{\mathbb{G}}} \mathrm{E}_0\Big[\sum_{t=0}^{T}(1-\widehat{\nu}_t(c_t))^t\Big\{F(c_t, \widehat{\nu}_t(c_t)) - \lambda_t(c_t)\Big\}\Big]\,,$$

and the penalties λ_t are chosen from the feasible set

$$\mathcal{L}_{\mathbb{F}} = \{\lambda \in \mathcal{L} : \mathrm{E}_0\big[\lambda(C,\omega)\big] \leq 0\,, \quad \forall C \in \mathcal{A}^{\mathbb{F}}\}\,.$$

Proof.

$$\begin{aligned}
V(x_0) &= \max_{C \in A^{\mathbb{F}}} \min_{\nu_t} \mathrm{E}_0\Big[\sum_{t=0}^{T}(1-\nu_t)^t F(c_t,\nu_t)\Big] \\
&= \max_{C \in A^{\mathbb{F}}} \mathrm{E}_0\Big[\sum_{t=0}^{T}\big(1-\nu_t^*(c_t)\big)^t F(c_t,\nu_t^*(c_t))\Big] \\
&\leq \max_{C \in A^{\mathbb{F}}} \mathrm{E}_0\Big[\sum_{t=0}^{T}(1-\widehat{\nu}_t(c_t))^t F(c_t,\widehat{\nu}_t(c_t))\Big] \\
&\leq \max_{C \in A^{\mathbb{F}}} \mathrm{E}_0\Big[\sum_{t=0}^{T}(1-\widehat{\nu}_t(c_t))^t\Big\{F(c_t,\widehat{\nu}_t(c_t)) - \lambda_t(c_t)\Big\}\Big] \\
&\leq \max_{C \in A^{\mathbb{G}}} \mathrm{E}_0\Big[\sum_{t=0}^{T}(1-\widehat{\nu}_t(c_t))^t\Big\{F(c_t,\widehat{\nu}_t(c_t)) - \lambda_t(c_t)\Big\}\Big] = V^{\mathcal{R}}(x_0)\,.
\end{aligned}$$

■

The first inequality arises from a sub-optimal choice of the discount rate process, the second because of the super-martingale property of λ_t, and the final one because $A^{\mathbb{F}} \subseteq A^{\mathbb{G}}$. The value function of the relaxed problem equals the value function of the original problem if $\widehat{\nu} = \nu^*$ and the penalty is chosen ideally: $\lambda_t^* = \mathrm{E}\,[V_t(x_{t+1})|\mathcal{G}_t] - \mathrm{E}\,[V_t(x_{t+1})|\mathcal{F}_t]$. The proof is similar to the one for separable preferences (see e.g. Brown et al. (2010)) and is omitted.

The expected utility from adoption of a heuristic policy $\widehat{c}_t(x_t)$ provides a lower bound on the agent's value function. This is computed recursively starting with the boundary condition $\widehat{V}_T(x_T) = u_T(c_T)$, and proceeding backwards.

C Transition matrix

The transition matrix that we use for Krusell-Smith's model in Section 3 is the same as used in den Haan, Judd, and Juillard (2010):

z, ϵ/z', ϵ'	$(z_b, 0)$	$(z_b, 1)$	$(z_g, 0)$	$(z_g, 1)$
$(z_b, 0)$	0.525000	0.350000	0.031250	0.093750
$(z_b, 1)$	0.038889	0.836111	0.002083	0.122917
$(z_g, 0)$	0.093750	0.031250	0.291667	0.583333
$(z_g, 1)$	0.009115	0.115885	0.024306	0.850694

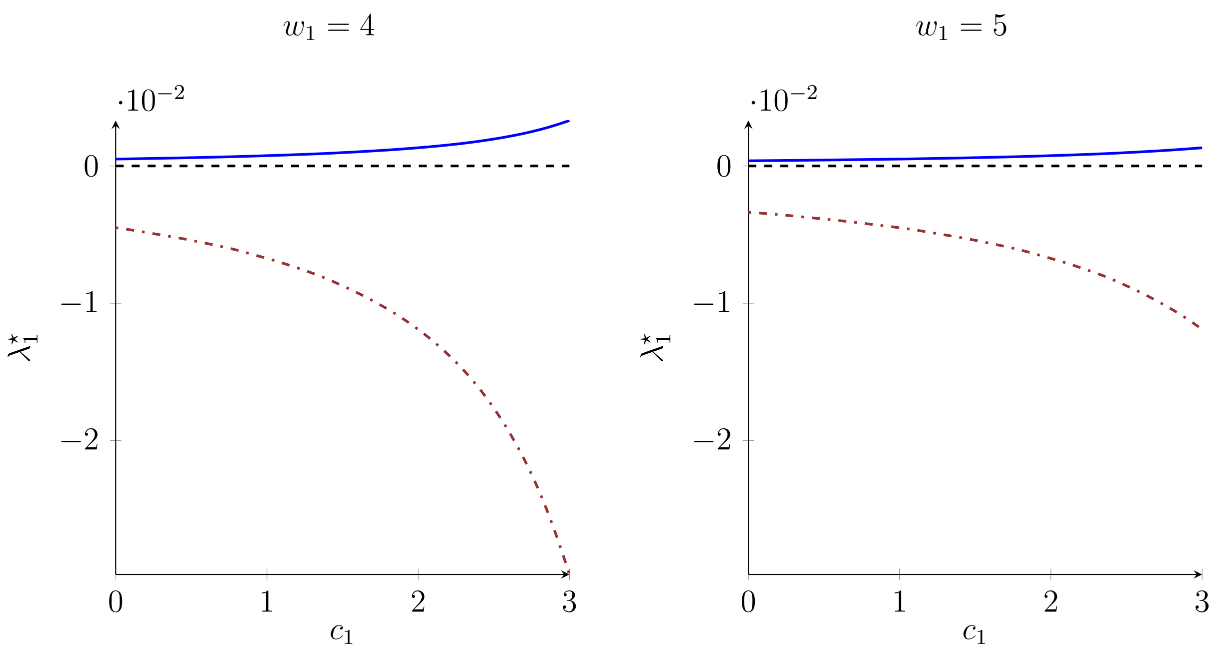

Figure 1: Ideal penalty $\lambda_1^\star$ for the consumption-saving problem of Section 2.1, plotted as a function of time-1 consumption choice c_1. The risk-free interest rate $R = 1.02$. Labor income follows a two-state Markov chain with $y_L = 1.0$, and $y_H = 4.0$ with independent draws each period. The probability of y_H is $p = 0.9$. The agent has power utility with relative risk aversion of 5, and time preference parameter $\beta = 0.9$. Time horizon is $T = 100$. The left panel corresponds to the time-0 wealth $w_0 = 4$, the right panel corresponds to $w_0 = 5$. The solid line plots the value of the penalty in state $y_2 = y_H$, while the dash-dot line plots $\lambda_1^\star$ for $y_2 = y_L$. The dash line shows the expectation of the penalty over the two possible realizations of y_2, which is identically equal to zero.

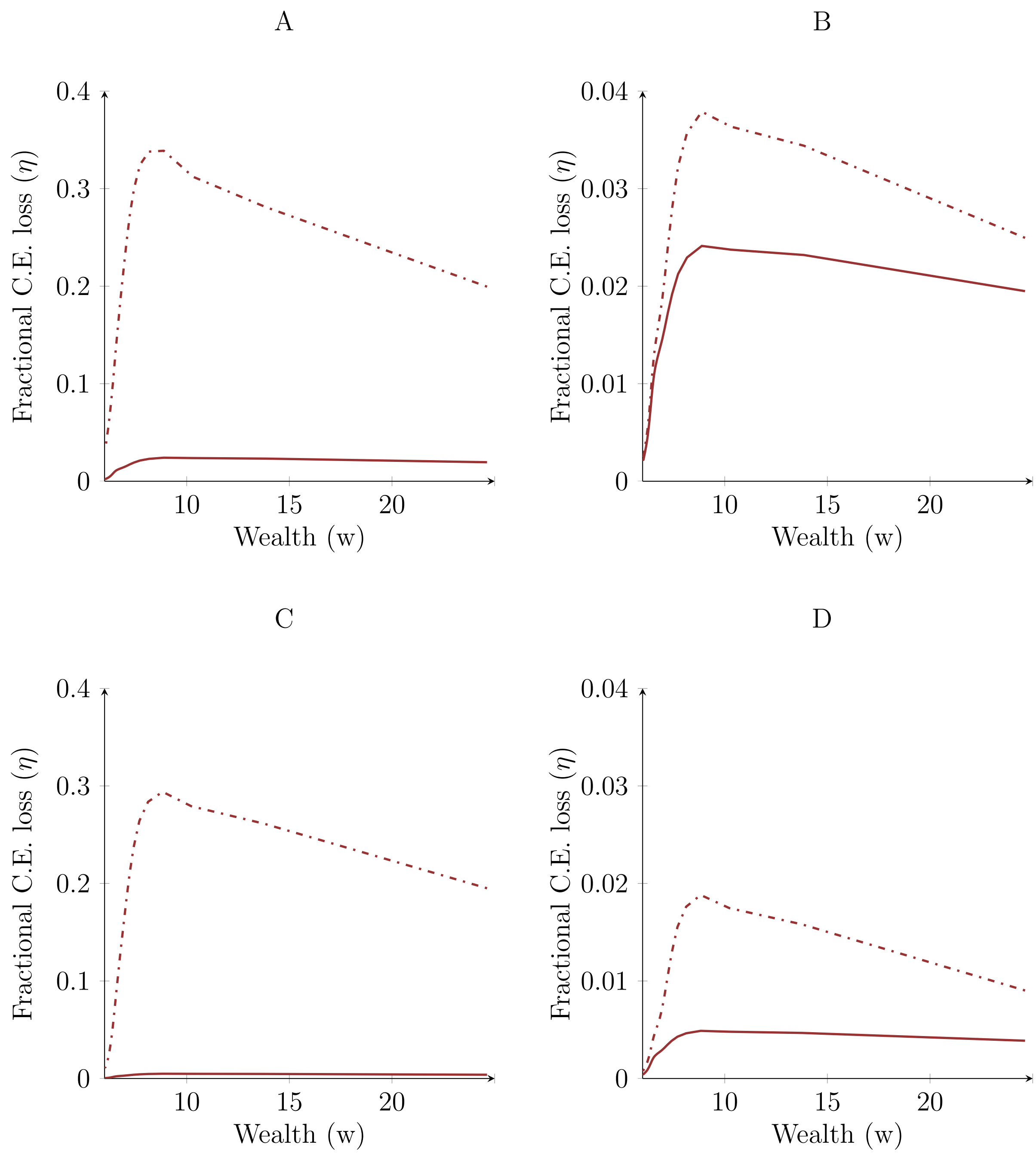

Figure 2: Fractional certainty equivalent loss as a function of wealth for the consumption-saving problem of Section 2.1. The risk-free interest rate $R = 1.02$. Labor income follows a two-state Markov chain with $y_L = 1.0$, and $y_H = 4.0$ with independent draws each period. The probability of y_H is $p = 0.9$. The agent has power utility with relative risk aversion of 5, and time preference parameter $\beta = 0.9$. Time horizon is $T = 100$. We use 500 paths to estimate the value function of the relaxed problem. In panels A and B, the agent uses a particular sub-optimal policy, which is optimal for $\widehat{p} = 0.895$, while in panels C and D he uses a policy optimal for $\widehat{p} = 0.899$. The dashed line in all panels show the upper bound, while the solid line shows the actual welfare loss. Panels A and C show the upper bound on welfare loss under zero penalty, while in panels B and D the upper bound is based on the value function computed using the corresponding sub-optimal policies.

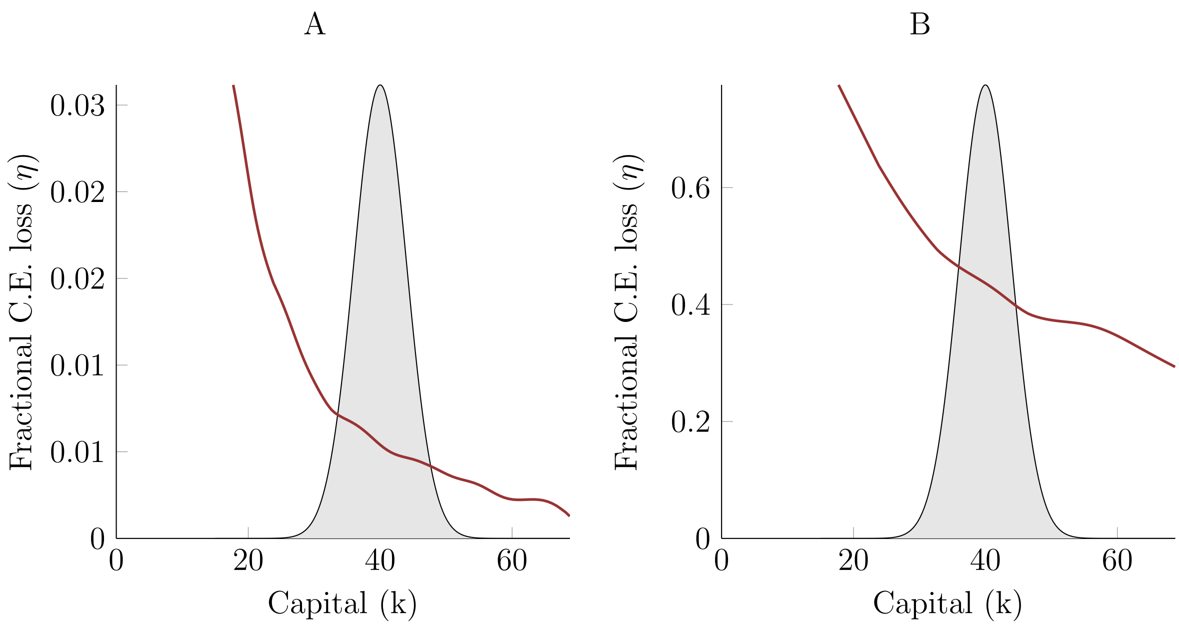

Figure 3: Upper bound on welfare loss for the incomplete markets model of Krusell and Smith where the approximate equilibrium is computed using moment truncation. All parameters values are identical to those in Krusell and Smith (1998). Fractional certainty equivalent loss is defined in (22). The value function of the relaxed problem is estimated by averaging over 100 paths of aggregate shocks. The agent is unemployed and the aggregate state of the economy is low. The cross-sectional distribution of capital shown by the area under the shaded curve is the stochastic steady-state distribution. Panel B shows the upper bound on welfare loss under zero penalty, while in panel A the upper bound is based on the approximate value function of the agent, according to (21).

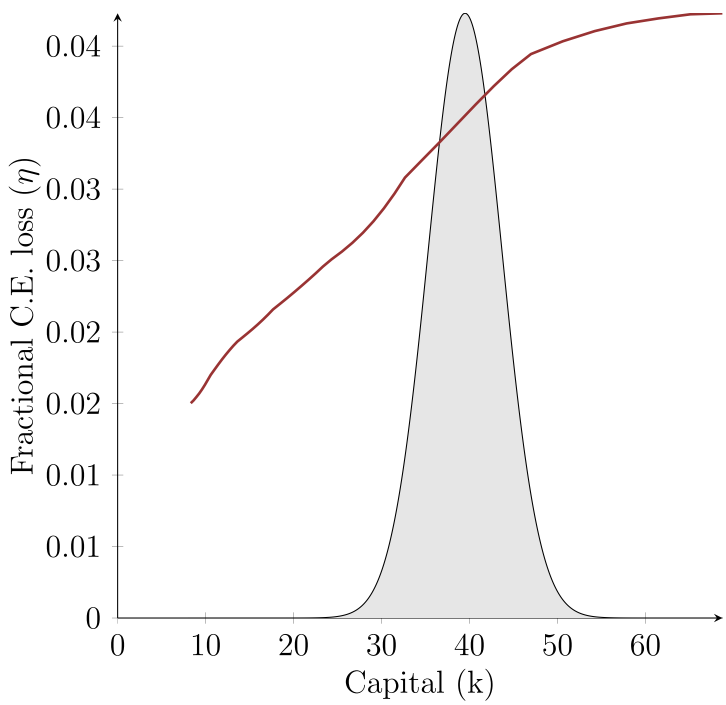

Figure 4: Upper bound on welfare loss for the incomplete markets model of Krusell and Smith where the approximate equilibrium is computed using history truncation. All parameters values are identical to those in Krusell and Smith (1998). Fractional certainty equivalent loss is defined in (22). The value function of the relaxed problem is estimated by averaging over 100 paths of aggregate shocks. The agent is unemployed and the aggregate state of the economy is low. The cross-sectional distribution of capital shown by the area under the shaded curve is the stochastic steady-state distribution.

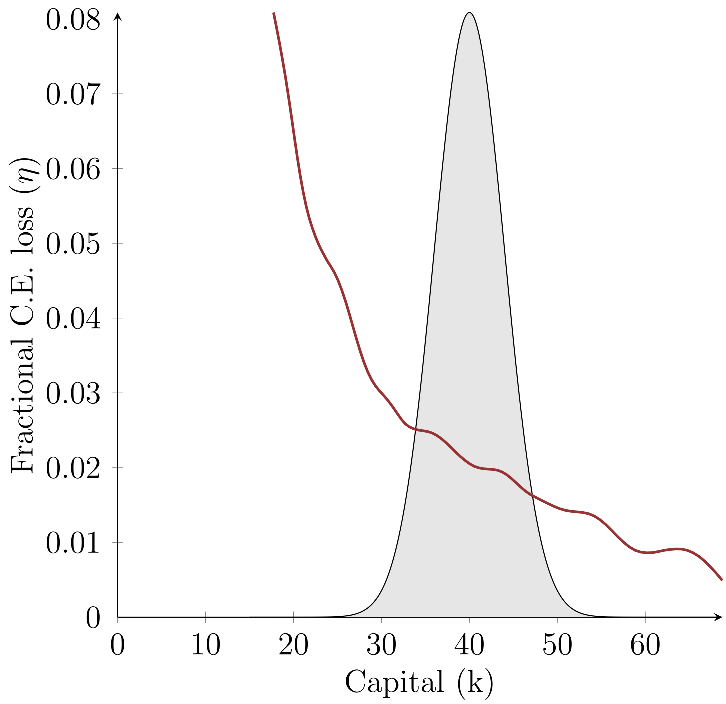

Figure 5: Upper bound on welfare loss for the incomplete markets model of Krusell and Smith where the approximate equilibrium is computed using moment truncation. Starting from the steady-state, aggregate productivity experiences a one-time five-fold permanent increase in volatility. We compute agent policies under the new parameter values. Agents immediately switch to new policies following the regime change. Fractional certainty equivalent loss is defined in (22). The value function of the relaxed problem is estimated by averaging over 100 paths of aggregate shocks. The agent is unemployed and the aggregate state of the economy is low. The area under the shaded curve shows the cross-sectional distribution of capital, and corresponds to the stochastic steady state of the economy before the permanent shock is realized.

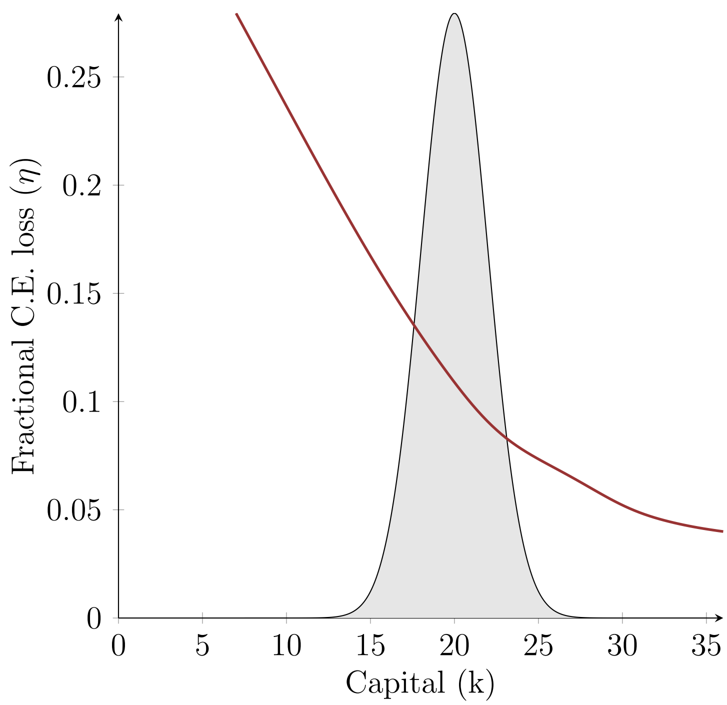

Figure 6: Upper bound on welfare loss for the transitional dynamics following a 50% loss in capital stock of all agents in the incomplete markets model of Krusell and Smith. The approximate equilibrium and policies are computed using moment truncation. Fractional certainty equivalent loss is defined in (22). The value function of the relaxed problem is estimated by averaging over 100 paths of aggregate shocks. The agent is unemployed and the aggregate state of the economy is low. The area under the shaded curve shows the cross-sectional distribution of capital immediately after the economy-wide capital loss is realized.

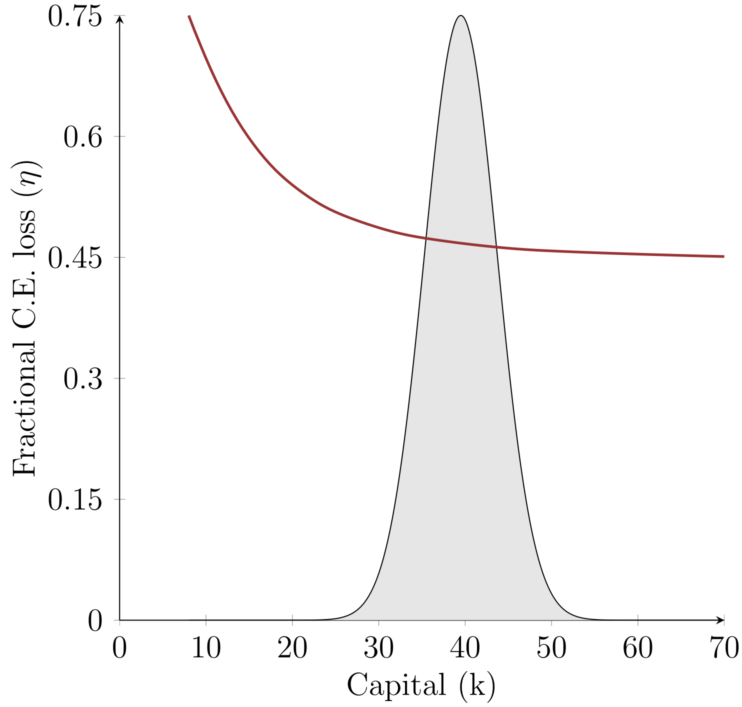

Figure 7: Upper bound on welfare loss for the incomplete markets model of Krusell and Smith where the approximate equilibrium is computed using history truncation. Starting from the steady-state, aggregate productivity experiences a one-time five-fold permanent increase in volatility. We compute agent policies under the new parameter values. Agents switch to new policies three periods after the regime change. Fractional certainty equivalent loss is defined in (22). The value function of the relaxed problem is estimated by averaging over 100 paths of aggregate shocks. The agent is unemployed and the aggregate state of the economy is low. The area under the shaded curve shows the cross-sectional distribution of capital, and corresponds to the stochastic steady state of the economy before the permanent shock is realized.

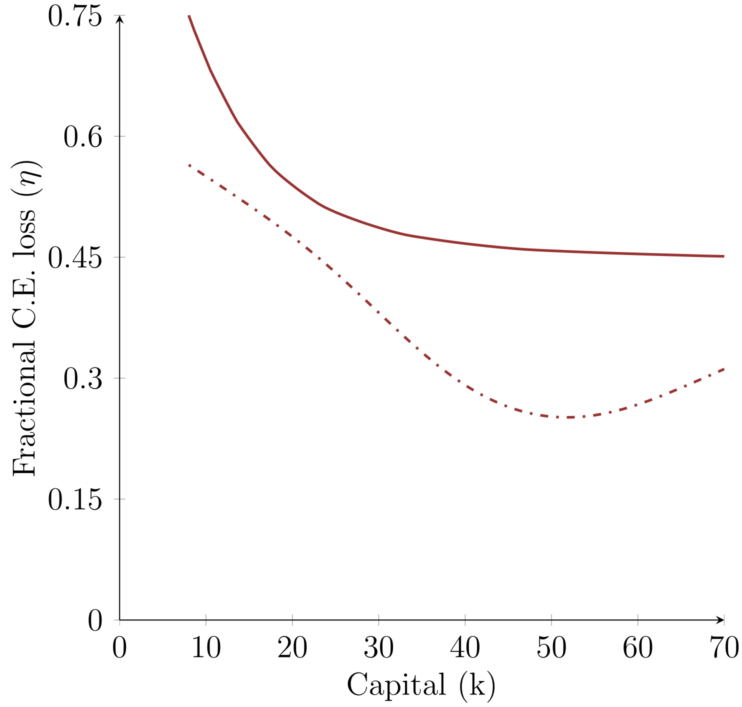

Figure 8: Approximate equilibrium in the incomplete markets model of Krusell and Smith, computed using history truncation, following a five-fold permanent increase in volatility of aggregate productivity. Fractional certainty equivalent loss is defined in (22). The dotted curve plots welfare loss relative to the feasible policy constructed using the Krusell-Smith moment truncation algorithm. The solid curve is the same curve as in Figure 7.

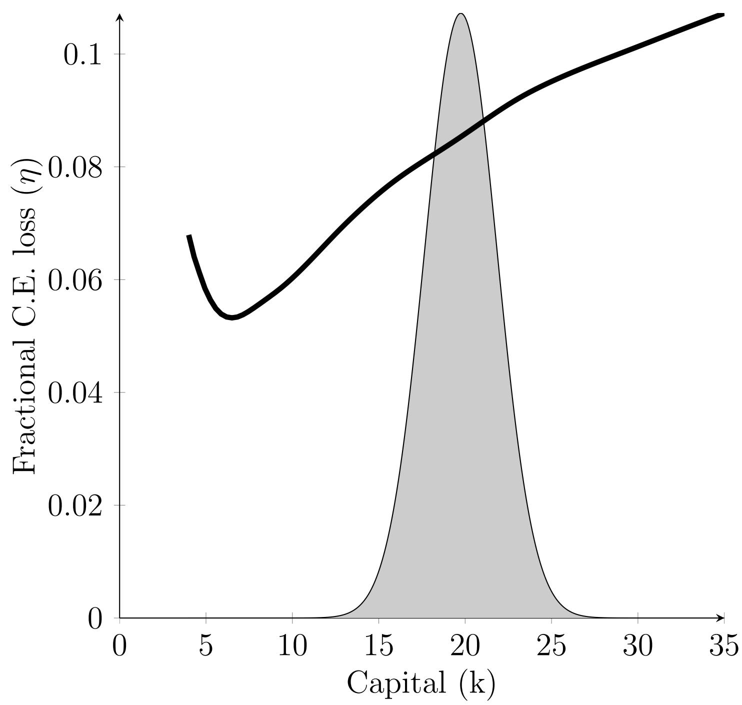

Figure 9: Upper bound on welfare loss for the transitional dynamics following a 50% loss in capital stock of all agents in the incomplete markets model of Krusell and Smith. The approximate equilibrium and policies are computed using history truncation. Fractional certainty equivalent loss is defined in (22). The value function of the relaxed problem is estimated by averaging over 100 paths of aggregate shocks. The agent is unemployed and the aggregate state of the economy is low. The area under the shaded curve shows the cross-sectional distribution of capital immediately after the economy-wide capital loss is realized.

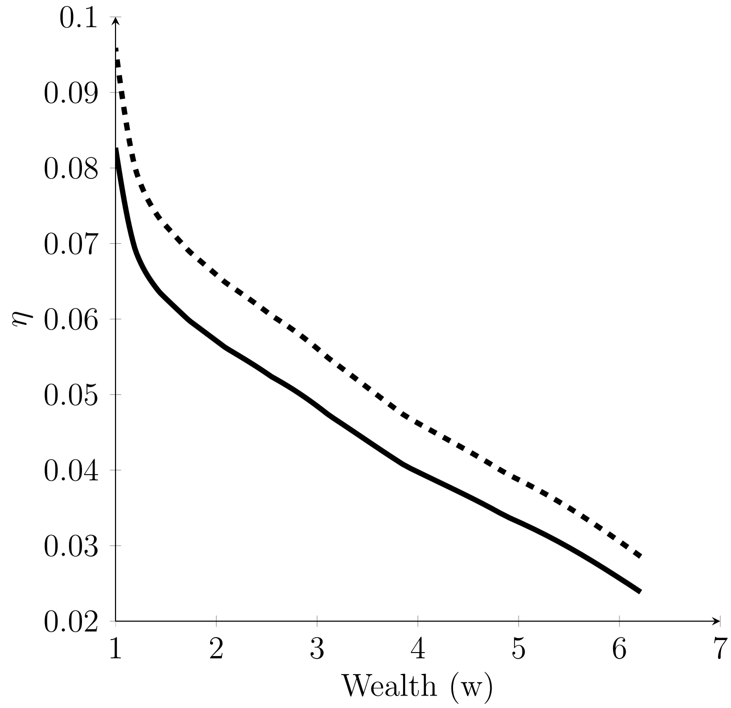

Figure 10: Fractional certainty equivalent loss as a function of wealth for an agent using sub-optimal policies to smooth labor income risk using a single risk-free asset with constant interest rate $R = 1.02$. Labor income follows a two-state Markov chain with $e_L = 1.0$, and $e_H = 4.0$ with independent draws each period. The probability of e_H is $p = 0.9$. The problem horizon is $T = 10$. The agent uses a policy that is optimal for $\widehat{p} = 0.895$. The dashed curve shows the upper bound, while the solid curve shows the actual welfare loss. The agent has Epstein-Zin utility with EIS equal to 1.5, relative risk aversion of 5, and time preference parameter $\beta = 0.9$. We use 500 simulated paths to estimated the value function of the relaxed problem.

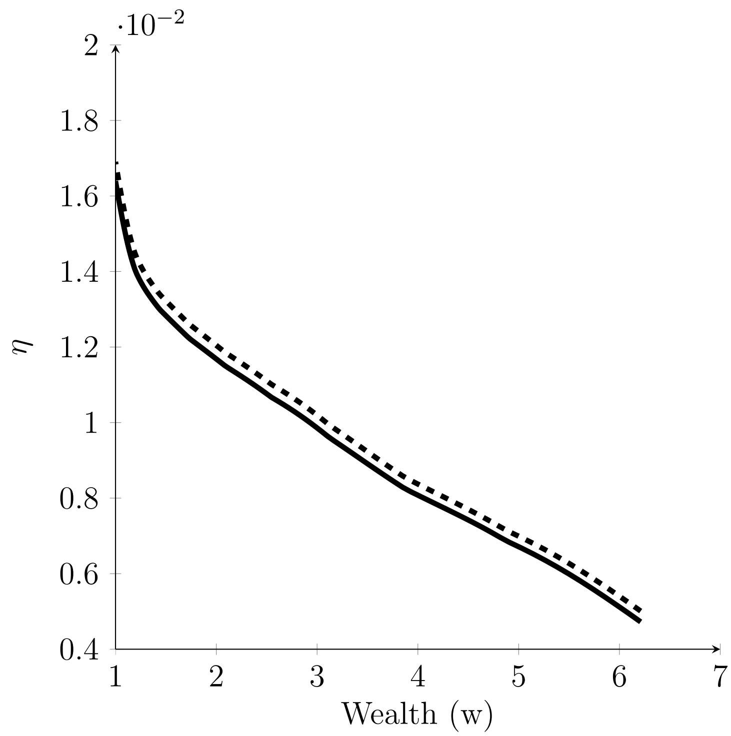

Figure 11: Fractional certainty equivalent loss as a function of wealth for the same model as in Figure 10. The agent uses a policy that is optimal for $\widehat{p} = 0.899$. The dashed curve shows the upper bound, while the solid curve shows the actual welfare loss. Parameters used are the same as in Figure 10.