Supplementary Appendix (Not for Publication)

B Supplementary Analytical Appendix

B.1 Two-Generations Model: Convexity of Welfare Gain

Proposition 3. *Applying Definition 1 to equation (4) gives the properties of the components of the CEV as follows:*

$$\frac{\partial dg_c^{PE}(i)}{\partial \theta} > 0, \textit{ for } i \in \{AR, IR\}, \qquad \frac{\partial \Delta_{CWG}}{\partial \theta} > 0, \qquad \frac{\partial \frac{\Delta_{CWG}}{dg_c^{PE}(AR)}}{\partial \theta} > 0.$$

Proof. The proof consists of three steps. First, we translate the CEV as a function of variances of random variables in logs into the respective terms in levels. Second, we combine the general definition (1) with the formula for the CEV in equation (4) to write the respective terms for $g_c(0,0), dg_c(AR)$ and so forth. Third, we derive the partial derivatives.

1. By log-normality we have $\exp\left(\theta\left(\sigma^2_{\ln\eta} + \sigma^2_{\ln AR}\right)\right) = \left((1+\sigma^2_\eta)(1+\sigma^2_{AR})\right)^\theta$, where $\sigma_{AR} \equiv \sqrt{\sigma^2_\zeta + \sigma^2_\varrho + \sigma^2_\zeta\sigma^2_\varrho}$ and therefore

$$g_c^{PE} = \left(\frac{1+\lambda}{\bar{R}}\left((1+\sigma^2_\eta)(1+\sigma^2_{AR})\right)^\theta - 1\right)d\tau \tag{17}$$

   which encompasses the case $\theta = 1$ shown in Section 2.2.

2. Applying definition (1) to (17) readily gives

$$\begin{aligned}
g_c^{PE}(AR, IR) &= \left(\frac{1+g}{\bar{R}}\left(\left(1+\sigma^2_\eta\right)\left(1+\sigma^2_{AR}\right)\right)^\theta - 1\right)d\tau \\
g_c^{PE}(0,0) &= \left(\frac{1+g}{\bar{R}} - 1\right)d\tau \\
g_c^{PE}(AR,0) &= \left(\frac{1+g}{\bar{R}}\left(1+\sigma^2_{AR}\right)^\theta - 1\right)d\tau \;\Leftrightarrow\; dg_c(AR) = \frac{1+g}{\bar{R}}\left(\left(1+\sigma^2_{AR}\right)^\theta - 1\right)d\tau \\
g_c^{PE}(0,IR) &= \left(\frac{1+g}{\bar{R}}\left(1+\sigma^2_\eta\right)^\theta - 1\right)d\tau \;\Leftrightarrow\; dg_c(IR) = \frac{1+g}{\bar{R}}\left(\left(1+\sigma^2_\eta\right)^\theta - 1\right)d\tau
\end{aligned}$$

   We observe that $dg_c(AR)$ and $dg_c(IR)$ are both increasing in θ. From these terms we further get

$$\begin{aligned}
\Delta_{CWG} &= g_c^{PE}(AR, IR) - \left(g_c^{PE}(0,0) + dg_c(AR) + dg_c(IR)\right) \\
&= \frac{1+g}{\bar{R}} \left(\left(\left(1+\sigma_\eta^2\right) \left(1+\sigma_{AR}^2\right) \right)^\theta - 1 \right. \\
&\quad \left. - \left(\left(1+\sigma_{AR}^2\right)^\theta - 1 \right) - \left(\left(1+\sigma_\eta^2\right)^\theta - 1 \right) \right) d\tau \\
|_{\theta=1} &= \frac{1+g}{\bar{R}} \sigma_\eta^2 \sigma_{AR}^2
\end{aligned}$$

and readily observe that Δ_{CWG} is increasing in σ_η as well as σ_{AR}.

3. To establish that Δ_{CWG} is also increasing in θ, simplify notation by defining $\sigma_{TR} = \sqrt{\sigma_\eta^2 + \sigma_{AR}^2 + \sigma_\eta^2\sigma_{AR}^2}$ where TR stands in for "total risk". Using this notation, observe that

$$\frac{\partial \Delta_{CWG}}{\partial \theta} = \frac{1+g}{\bar{R}} \Big(\ln(1+\sigma_{TR}^2)(1+\sigma_{TR}^2)^\theta - \ln(1+\sigma_{AR}^2)(1+\sigma_{AR}^2)^\theta - \ln(1+\sigma_\eta^2)(1+\sigma_\eta^2)^\theta \Big)$$

Evaluate this at $\theta = 1$ to get

$$\left.\frac{\partial \Delta_{CWG}}{\partial \theta}\right|_{\theta=1} = \frac{1+g}{\bar{R}} \left(\sigma_{AR}^2 \cdot \ln\left(\frac{1+\sigma_{TR}^2}{1+\sigma_{AR}^2}\right) + \sigma_\eta^2 \cdot \ln\left(\frac{1+\sigma_{TR}^2}{1+\sigma_\eta^2}\right) + \ln(1+\sigma_{TR}^2) \cdot \sigma_{AR}^2 \cdot \sigma_\eta^2 \right) > 0.$$

The general conclusion that $\frac{\partial \Delta_{CWG}}{\partial \theta} > 0$ for all θ then follows from continuity. Finally, we can express the contribution to the CEV of CWG relative to AR as

$$\frac{d\Delta_{CWG}}{dg_c(\sigma_{AR})} = \frac{\left(\left(1+\sigma_\eta^2\right)(1+\sigma_{AR}^2)\right)^\theta}{(1+\sigma_{AR}^2)^\theta - 1} - \frac{1}{(1+\sigma_{AR}^2)^\theta - 1} - 1 - \frac{\left(1+\sigma_\eta^2\right)^\theta - 1}{(1+\sigma_{AR}^2)^\theta - 1}$$

Take the derivative of this term w.r.t. θ to get

$$\frac{\partial \frac{d\Delta_{CWG}}{dg_c(\sigma_{AR})}}{\partial \theta} = \frac{1}{\left((1+\sigma_{AR}^2)^\theta - 1\right)^2} \left(\ln\left(\left(1+\sigma_\eta^2\right)\left(1+\sigma_{AR}^2\right)\right)\left(\left(1+\sigma_\eta^2\right)\left(1+\sigma_{AR}^2\right)\right)^\theta \right.$$
$$\left(\left(1+\sigma_{AR}^2\right)^\theta - 1\right) - \left(\left(1+\sigma_\eta^2\right)\left(1+\sigma_{AR}^2\right)\right)^\theta \ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right)^\theta$$
$$+ \ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right)^\theta$$
$$- \ln\left(1+\sigma_\eta^2\right)\left(1+\sigma_\eta^2\right)^\theta \left(\left(1+\sigma_{AR}^2\right)^\theta - 1\right)$$
$$\left. + \left(\left(1+\sigma_\eta^2\right)^\theta - 1\right) \ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right)^\theta \right).$$

Evaluated at $\theta = 1$ we get

$$\left.\frac{\partial \frac{\Delta_{CWG}}{dg_c(\sigma_{AR})}}{\partial \theta}\right|_{\theta=1} = \frac{1}{(\sigma_{AR}^2)^2} \left(\ln\left(\left(1+\sigma_\eta^2\right)\left(1+\sigma_{AR}^2\right)\right)\left(1+\sigma_\eta^2\right)\left(1+\sigma_{AR}^2\right)\sigma_{AR}^2 \right.$$
$$- \left(1+\sigma_\eta^2\right)\left(1+\sigma_{AR}^2\right)\ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right) + \ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right)$$
$$\left. - \ln\left(1+\sigma_\eta^2\right)\left(1+\sigma_\eta^2\right)\sigma_{AR}^2 + \sigma_\eta^2 \ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right) \right)$$

Now split the numerator up as follows:

$$N \equiv \underbrace{\ln\left(1+\sigma_{TR}^2\right)\left(1+\sigma_{TR}^2\right)\sigma_{AR}^2 - \left(1+\sigma_{TR}^2\right)\ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right)}_{\equiv \Psi_1}$$
$$+ \underbrace{\ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right) - \ln\left(1+\sigma_\eta^2\right)\left(1+\sigma_\eta^2\right)\sigma_{AR}^2 + \sigma_\eta^2 \ln\left(1+\sigma_{AR}^2\right)\left(1+\sigma_{AR}^2\right)}_{\equiv \Psi_2}$$

where σ_{TR}^2 is again the variance due to total risk. Next, notice that:

$$\Psi_1 = (1+\sigma_{TR}^2)\left(\sigma_{AR}^2 \ln(1+\sigma_\eta^2) - \ln(1+\sigma_{AR}^2)\right)$$
$$\Psi_2 = (1+\sigma_{TR}^2)\ln\left(1+\sigma_{AR}^2\right) - \sigma_{AR}^2\left(1+\sigma_\eta^2\right)\ln\left(1+\sigma_\eta^2\right)$$
$$\Psi_1 + \Psi_2 = \left(\sigma_{AR}^2\right)^2 \ln(1+\sigma_\eta^2)\left(1+\sigma_\eta^2\right).$$

Therefore $\left.\frac{\partial \frac{\Delta_{CWG}}{dg_c(\sigma_{AR})}}{\partial \theta}\right|_{\theta=1} > 0$ and the conclusion for general θ again follows by continuity.

□

B.2 General Equilibrium Extension of the Simple Model

We sketch the main elements, provide and extend the key findings of Harenberg and Ludwig (2015) with the purpose (i) to show that the discount rate plays an additional decisive role for evaluating the welfare effects of social security and (ii) to expose analytically one channel for the ambiguity of risk interactions on the welfare consequences from crowding out of capital. To this end, consider a two period extension of the model of Section 2 in general equilibrium. In the first period, households are endowed with one unit of labor productivity and work full time. In the second period, they only work fraction $\omega \in [0, 1)$ of their time and are retired with fraction $1-\omega$. We also assume that the idiosyncratic shock only hits in the second period of life. The subperiod structure together with the assumption that the shock only hits in the second period enables us to model precautionary savings behavior while maintaining closed form solutions in general equilibrium when we focus on a logarithmic utility function and Cobb-Douglas production.[45]

B.2.1 Modifications

The modifications of the simple model of Section 2 are threefold. First, denoting the discount factor by β, expected life-time utility is $E_t\left[\ln(c_{1,t}) + \beta \ln(c_{2,t+1})\right]$, where we restrict the analysis to log utility for analytical solutions. Second, the budget constraints in the two periods now write as

$$s_{2,t+1} + c_{1,t} = (1-\tau_t)w_t \text{ and } c_{i,2,t+1} \leq s_{2,t+1}R_{t+1} + \omega(1-\tau)w_{t+1}\eta_{i,2,t+1} + (1-\omega)y^{ss}_{t+1}.$$

Third, to close the model in general equilibrium, production takes place with a representative firm's production function $F(K_t, \Upsilon_t L_t)$, where K_t is aggregate capital, $L_t = N_{t,0} + \omega N_{t,1}$ is aggregate labor, and Υ_t is labor augmenting technological progress, growing at exogenous rate λ. Next, introduce shock ζ_t as a standard RBC shock to output and shock ϱ_t as a shock to the user costs of capital and assume 100% depreciation (again for analytical reasons) so that profits are given by $\Pi_t = \zeta_t F(K_t, \Upsilon_t L_t) - (1+r_t)\,\varrho_t^{-1} K_t - w_t L_t$. Denoting by $k_t = \frac{K_t}{\Upsilon_t L_t}$ the capital stock per efficiency unit (the capital "intensity"), profit maximization of this neoclassical firm

[45] Because of the human capital wealth effect from second period income, we cannot derive closed form solutions even with logarithmic utility in partial equilibrium. Analytical tractability only arises when plugging in the general equilibrium dynamics into the first-order conditions of households, also see Krueger and Ludwig (2007), Ludwig and Vogel (2009), and Krueger and Ludwig (2017). Formally, the reason is that second period income is discounted with the market interest rate and both income and the interest rate are functions of capital in general equilibrium.

then gives the first order conditions as

$$w_t = (1-\alpha)\Upsilon_t k_t^{\alpha}\zeta_t = \bar{w}_t\zeta_t \quad \text{and} \quad R_t = \alpha k_t^{\alpha-1}\zeta_t\varrho_t = \bar{R}_t\zeta_t\varrho_t.$$

which is the general equilibrium analogue to equation (2).

B.2.2 Analysis

In general equilibrium, the law of motion of the capital intensity writes as

$$k_{t+1} = \frac{1}{(1+\lambda)(1+\omega)} s(\tau)(1-\tau)(1-\alpha)\zeta_t k_t^{\alpha} \tag{18}$$

$$\text{where} \quad s(\tau) = \frac{\beta\Gamma(\tau)}{1+\beta\Gamma(\tau)} \leq \frac{\beta}{1+\beta} \tag{19}$$

$$\text{and} \quad \Gamma(\tau) = \mathbb{E}_t\left[\frac{1}{1+\frac{1-\alpha}{\alpha(1+\omega)\varrho_{t+1}}\left(\omega\eta_{i,2,t+1}+\tau\left(1+\omega(1-\eta_{i,2,t+1})\right)\right)}\right] \leq 1. \tag{20}$$

cf. Propostion 3 in Harenberg and Ludwig (2015). To interpret this, notice that in a deterministic economy (where $\zeta_t = \varrho_t = \eta_t = 1$) without work effort in the second period ($\omega = 0$) equation (18) is just the standard textbook variant of the law of motion of the capital intensity in a 2-period OLG economy with logarithmic utility and Cobb-Douglas production. The risk adjustment term $\Gamma(\tau)$ in equation (19) captures (i) precautionary saving behavior w.r.t. idiosyncratic income risk—$s(\tau)$ increases in response to a mean-preserving spread of η; (ii) intertemporal reallocation w.r.t. return risk—$s(\tau)$ decreases in response to a mean-preserving spread of ϱ because savings become less attractive if the return risk increases; and (iii) crowding out of savings—the saving rate decreases when τ is increased. In the deterministic economy, life-cycle savings are reduced due to positive retirement income in the second period. In the stochastic economy, there is additional partial insurance of idiosyncratic risk through social security which decreases precautionary savings.

Based on this structure, Proposition 4 in Harenberg and Ludwig (2015) contains the main results on the welfare benefits from insurance and the welfare costs from crowding out in terms of utility units. Our next proposition extends those results to a consumption equivalent variation:

Proposition 4. *The CEV from a marginal introduction of social security at rate $d\tau > 0$ in the* *stationary equilibrium is given by* $g_c^{GE} = g_c^{PE} + g_c^{CO}$*, where*

$$g_c^{PE} \approx \frac{1}{1+\beta}\left(\beta\mathbb{E}\left[\frac{\frac{1-\alpha}{\alpha}\frac{1}{\varrho_{t+1}} - \frac{1-\alpha}{\alpha}\frac{\omega}{1+\omega}\frac{\eta_{i,2,t+1}}{\varrho_{t+1}} - 1}{1 + \frac{(1-\alpha)}{\alpha}\frac{\omega}{1+\omega}\frac{\eta_{i,2,t+1}}{\varrho_{t+1}}}\right] - 1\right) \tag{21}$$

$$g_c^{CO} \approx -\frac{1}{1+\beta}\left(\left(1 - \varphi_{s,\tau}|_{\tau=0}\right)\left(\frac{\alpha}{1-\alpha}\frac{1+\beta}{\beta} - \Gamma|_{\tau=0}\right)\right). \tag{22}$$

$\Gamma|_{\tau=0}$ *is term (20) for* $\tau = 0$ *and* $\varphi_{s,\tau}|_{\tau=0}$ *is the semi-elasticity of the saving rate w.r.t.* τ*,* $\varphi_{s,\tau} = \frac{\partial s/s}{\partial \tau}$ *, again evaluated at* $\tau = 0$*.*

Proof of Proposition 4.

1. Harenberg and Ludwig (2015), Proposition 4, show that social security increases ex-ante welfare in the stationary equilibrium if and only if

$$A + B > 0$$

$$\text{where} \quad A \equiv \beta\mathbb{E}\left[\frac{\frac{1-\alpha}{\alpha}\frac{1}{\varrho_{t+1}} - \frac{1-\alpha}{\alpha}\frac{\omega}{1+\omega}\frac{\eta_{i,2,t+1}}{\varrho_{t+1}} - 1}{1 + \frac{(1-\alpha)}{\alpha}\frac{\omega}{1+\omega}\frac{\eta_{i,2,t+1}}{\varrho_{t+1}}}\right] - 1$$

$$\text{and} \quad B \equiv -\beta\left(1 - \varphi_{s,\tau}|_{\tau=0}\right)\left(\frac{\alpha}{1-\alpha}\frac{1+\beta}{\beta} - \Gamma|_{\tau=0}\right).$$

2. To translate this into a CEV, observe that $EU\left(C^{\tau>0}\right) = EU\left(C^{\tau=0}\right) + (A+B)\,d\tau$, where capital letter U denotes life-time utility and capital letter C the consumption stream under the respective policy, hence $U\left(C^{\tau=0}\right) = \ln(c_{t,1}) + \beta\ln(c_{i,t+1,2})$. Using the definition of the CEV it is then straightforward to show that the CEV is given by $g_c = \exp\left(\left(\frac{A+B}{1+\beta}\right)d\tau\right) - 1$, which can be approximated as $g_c \approx \frac{A+B}{1+\beta}d\tau$.

□

Observe that g_c^{PE} is the CEV in "partial equilibrium", corresponding to our definition in the quantitative model of Section 3. It is the consumption equivalent variation from a marginal introduction of social security where the price sequence is held constant (i.e., does not react to social security) at the corresponding price sequence in the general equilibrium of the $\tau = 0$ economy. g_c^{CO} is the additional effect from the crowding out of capital induced by the social security reform.[46]

To interpret terms g_c^{PE} and g_c^{CO} further, we first restate as observation important findings from Harenberg and Ludwig (2015) and then move on to a Corollary:

[46]Also notice that the aggregate productivity shock ζ_t does not show up in (21) and (22). This is a consequence of (i) ζ_t showing up in all sources of income—wages, interest rates, and social security—, (ii) ζ_t showing up multiplicatively in the law of motion of the aggregate capital stock, and (iii) ζ_t also showing up multiplicatively in the derivatives of the utility function by the assumption of log utility. All these cancel each other out.

Observation 1. *1. The deterministic $\omega = 0$ economy is dynamically efficient if and only if*

$$\frac{\alpha}{1+\alpha}\frac{1+\beta}{\beta} > 1. \tag{23}$$

2. *If condition (23) holds then $g_c^{GE} < 0$ in the corresponding stochastic economy with $0 \leq \omega < 1$.*

3. *With respect to term g_c^{PE} we find that $\frac{\partial g_c^{PE}}{\partial \sigma_\varrho^2} > 0$, $\frac{\partial g_c^{PE}}{\partial \sigma_\eta^2} > 0$, $\frac{\partial^2 g_c^{PE}}{\partial \sigma_\varrho^2 \partial \sigma_\eta^2} > 0$.*

4. *With respect to term g_c^{CO} we find that $\frac{\partial \Gamma(\tau)|_{\tau=0}}{\partial \sigma_\eta^2} < 0$, $\frac{\partial^2 \Gamma(\tau)|_{\tau=0}}{\partial \varrho^2 \partial \sigma_\eta^2} < 0$ and, under condition (23), $\frac{\partial g_c^{CO}}{\partial \sigma_\varrho^2} < 0$, $\frac{\partial g_c^{CO}}{\partial \sigma_\eta^2} \lesseqgtr 0$, $\frac{\partial^2 g_c^{CO}}{\partial \sigma_\varrho^2 \partial \sigma_\eta^2} \lesseqgtr 0$.*

Observation 1.3 confirms our findings from Section 2. Importantly, Observation 1.4 states that the direction of the change of the welfare costs from crowding out when idiosyncratic risk increases (and its interaction with aggregate risk) is ambiguous. To understand this finding, suppose first that crowding out increases in risk, i.e., that $\varphi_{s,\tau}|_{\tau=0}$ becomes more strongly negative as risk goes up.[47] This means that precautionary savings decrease more strongly if the amount of risk insured increases. Under this assumption, the effects of an increase of idiosyncratic risk on the costs of crowding out (as well as the cross partial w.r.t. aggregate return risk) are ambiguous. The formal reason for this finding is that $\Gamma|_{\tau=0}$ increases in idiosyncratic risk, so that the welfare costs of crowding out are less strong. Intuitively, while crowding out reduces the mean capital stock thereby leading to welfare losses (as in a deterministic dynamically efficient economy) a lower capital stock also reduces the exposure to idiosyncratic wage risk because wages depend positively on the capital stock and are multiplicative in the shock, an effect which is (at least partially) offsetting the utility consequences of crowding out.

The next corollary on the importance of the discount factor—which is not contained in Harenberg and Ludwig (2015)—provides further guidance for the calibration of our quantitative model and the interpretation of its results:

Corollary 1. *The CEVs from a marginal introduction of social security in partial equilibrium, g_c^{PE}, and, under condition (23), also in general equilibrium, g_c^{GE}, increase in β.*

[47] It is not possible to show this analytically, but it is the most plausible case. It is also found to hold in the numerical analysis in Harenberg and Ludwig (2015).

Proof of Corollary 1. Rewrite g_c^{PE} and g_c^{CO} as

$$g_c^{PE} \approx \frac{1}{1+\frac{1}{\beta}} \mathbb{E}\left[\frac{\frac{1-\alpha}{\alpha}\frac{1}{\varrho_{t+1}} - \frac{1-\alpha}{\alpha}\frac{\omega}{1+\omega}\frac{\eta_{i,2,t+1}}{\varrho_{t+1}} - 1}{1 + \frac{(1-\alpha)}{\alpha}\frac{\omega}{1+\omega}\frac{\eta_{i,2,t+1}}{\varrho_{t+1}}} \right] - \frac{1}{1+\beta}$$

$$g_c^{CO} \approx -\left(1 - \varphi_{s,\tau}|_{\tau=0}\right)\left(\frac{\alpha}{1-\alpha} - \frac{1}{1+\frac{1}{\beta}}\,\Gamma|_{\tau=0}\right)$$

and the result immediately follows. □

The intuition for the effect of discounting on the g_c^{PE} is that households value insurance of second period consumption more when β is increased. As to the intuition for g_c^{CO}, observe that increasing β lowers the welfare costs from crowding out—i.e., the distance $\frac{\alpha}{1-\alpha} - \frac{1}{1+\frac{1}{\beta}}\,\Gamma|_{\tau=0} > 0$ decreases towards zero—, because increasing β increases life-cycle savings which moves the economy closer to the boundary of dynamic inefficiency.

B.3 Definition of Recursive Markov Equilibrium

We here provide a formal definition of a competitive recursive Markov equilibrium, cf. Section 3.5. To this end, we define a state space that is sufficient for solving the households' maximization problem. Let $\mathcal{E} = \{e_1, e_2, ..., e_{max}\}$ and $\mathcal{J} = \{1, 2, ...J\}$, and let $\mathcal{M}$ be a sigma-algebra over $\{[\underline{\tilde{s}}, \overline{\tilde{s}}] \times [\underline{\tilde{b}}, \overline{\tilde{b}}] \times \mathcal{E} \times \mathcal{J}\}$, where $\underline{\tilde{s}}$, $\overline{\tilde{s}}$, $\underline{\tilde{b}}$, and $\overline{\tilde{b}}$ are the infimum and supremum on stock and bond holdings.[48] The measure Φ is defined over $\mathcal{M}$, and the set of all such measures is denoted by $\mathcal{Q}$. We follow the applied literature and define the state space to consist of Φ, the current idiosyncratic state $(\tilde{s}, \tilde{b}, e)$, and the current aggregate shock z. As a recursive equilibrium does not depend on the date-event, we drop time index t and use a prime for next period's variables. Finally, note that the economic dependency ratio, $p = \frac{P(z^t)}{L(z^t)} = \frac{\sum_{j=j_r}^{J}(1+n)^{J-j}}{\sum_{j=1}^{j_r-1}(1+n)^{J-j}\epsilon_j}$, and the labor-to-population ratio, $\ell = \frac{L(z^t)}{N(z^t)} = \frac{\sum_{j=1}^{j_r-1}(1+n)^{J-j}\epsilon_j}{\sum_{j=1}^{J}(1+n)^{J-j}}$, are both constant over time.

Definition 4. *For any initial $(z_0, \Phi_0) \in \mathcal{Z} \times \mathcal{Q}$, a recursive competitive equilibrium consists of a measure Φ, measurable functions for household choices $\{\tilde{c}_j(\tilde{s}, \tilde{b}, e; \Phi, z), \tilde{s}'_j(\tilde{s}, \tilde{b}, e; \Phi, z), \tilde{b}'_j(\tilde{s}, \tilde{b}, e; \Phi, z)\}$ and an associated value function $\tilde{v}_j(\tilde{s}, \tilde{b}, e; \Phi, z)$, firm choices $k(\Phi, z)$, social security settings $\{\tau, \tilde{y}^{ss}(\Phi, z)\}$, factor prices $\{\tilde{w}(\Phi, z), r(\Phi, z)\}$, asset returns $\{r_b(\Phi), r_s(\Phi, z)\}$, and a law of motion $H(\Phi, z, z')$ such that:*

[48] For a given level of aggregate capital and a given equity premium, the infimum and supremum on bond and stock holdings are implied by the bounds on the income process and the fact that households can't hold negative positions in the asset when they die, see Section 3.2. In equilibrium, aggregate capital and the equity premium will be bounded, and the infimum and supremum can be calculated for those bounded intervals.

a) *given functions for prices and returns and the law of motion, the households' policy functions* $\{\tilde{c}_j(\tilde{s},\tilde{b},e;\Phi,z),\tilde{s}'_j(\tilde{s},\tilde{b},e;\Phi,z),\tilde{b}'_j(\tilde{s},\tilde{b},e;\Phi,z)\}$ *solve*

$$\tilde{v}_j(\tilde{s},\tilde{b},e;\Phi,z)$$
$$=\max_{\tilde{c}>0,\tilde{s}',\tilde{b}'}\begin{cases}\left(\tilde{c}^{\frac{1-\theta}{\gamma}}+\tilde{\beta}\left(\sum_{z'}\sum_{e'}\pi_z(z'|z)\pi_e(e'|e)\,\tilde{v}_{j+1}^{1-\theta}\left(\tilde{s}',\tilde{b}',e';H(\Phi,z),z'\right)\right)^{\frac{1}{\gamma}}\right)^{\frac{\gamma}{1-\theta}} & \\ \tilde{c} & \text{if } j=J\end{cases}$$

$$\begin{aligned} \textit{s. t.} \qquad \tilde{c}+\tilde{s}'+\tilde{b}' &= (1+r_s(\Phi,z))\tilde{s}+(1+r_b(\Phi))\tilde{b} \\ &\quad +(1-\tau)\tilde{y}_j(e,\Phi,z)I(j)+\tilde{y}^{ss}(\Phi,z)(1-I(j))\ , \\ \tilde{y}_j(e,\Phi,z) &= \tilde{w}(\Phi,z)\epsilon_j\eta(e,z)\ , \\ \tilde{s}'+\tilde{b}' &\geq 0 \qquad\qquad \text{if } j=J, \end{aligned} \tag{24}$$

where $\tilde{\beta}=\beta(1+\lambda)^{\frac{1-\theta}{\gamma}}$*,*

b) *functions for prices and for firm choices are related by*

$$\begin{aligned}\tilde{w}(\Phi,z) &= (1-\alpha)\zeta(z)k(\Phi,z)^{\alpha},\\ r(\Phi,z) &= \alpha\zeta(z)k(\Phi,z)^{\alpha-1}-\delta(z),\end{aligned}$$

c) *functions for asset returns are given by*

$$\begin{aligned}r_b(\Phi) &= \frac{1}{\overline{\kappa}_f}E\left[r(\Phi,z)(1+\overline{\kappa}_f)-r_s(\Phi,z)\right],\\ r_s(\Phi,z) &= r(\Phi,z)(1+\overline{\kappa}_f)-\overline{\kappa}_f r_b(\Phi),\end{aligned}$$

d) *the pension system budget constraint holds, i.e.,*

$$\tau\tilde{w}(\Phi,z)=\tilde{y}^{ss}(\Phi,z)p, \tag{25}$$

where p *is the economic dependency ratio defined above,* *e) all markets clear:*

$$\zeta(z)k(\Phi,z)^{\alpha}+(1-\delta(z))k(\Phi,z)=\frac{1}{\ell}\sum_{j=1}^{J}\sum_{e}\int_{\tilde{b}}\int_{\tilde{s}}\tilde{c}_j(\tilde{s},\tilde{b},e;\Phi,z)\Phi(\tilde{s},\tilde{b},e,j)\,d\tilde{b}\,d\tilde{s}$$
$$+\,k(H(\Phi,z,z'),z')(1+\lambda)(1+n),$$
$$k(H(\Phi,z,z'),z')(1+\lambda)(1+n)=\frac{1}{\ell}\sum_{j=1}^{J}\sum_{e}\int_{\tilde{b}}\int_{\tilde{s}}(\tilde{s}'_j(\tilde{s},\tilde{b},e;\Phi,z)$$
$$+\,\tilde{b}'_j(\tilde{s},\tilde{b},e;\Phi,z))\Phi(\tilde{s},\tilde{b},e,j)\,d\tilde{b}\,d\tilde{s},$$
$$\frac{k(H(\Phi,z,z'),z')(1+\lambda)(1+n)}{(1+\overline{\kappa}_f)}=\frac{1}{\ell}\sum_{j=1}^{J}\sum_{e}\int_{\tilde{b}}\int_{\tilde{s}}\tilde{s}'_j(\tilde{s},\tilde{b},e;\Phi,z)\Phi(\tilde{s},\tilde{b},e,j)\,d\tilde{b}\,d\tilde{s},$$

and by Walras' Law, the bond market also clears,

f) the law of motion H is generated by the policy functions and the Markov transition matrix π_e, so that

$$\Phi'=H(\Phi,z,z')$$

with the initialization at $j=1$ of $\tilde{s}=\tilde{b}=0$.

B.4 Corollary: CEV in a Deterministic Economy, $g_c^{PE}(0,0)$

For an economy with an arbitrary number of generations J, we can provide a closed-form solution for g_c for an economy without risk. Following the discussion in Section 2.2, we denote the consumption equivalent variation in an economy without risk by $g_c^{PE}(0,0)$.

Corollary 2. *Denote by pvi_1^A (pvi_1^B) the present discounted value of lifetime income in policy A (B). The consumption equivalent variation in the partial equilibrium of the risk-free economy is given by*

$$g_c^{PE}(0,0)=\frac{\tilde{u}_1(\tilde{c}^B)}{\tilde{u}_1(\tilde{c}^A)}-1=\frac{\widetilde{pvi}_1^B}{\widetilde{pvi}_1^A}-1,$$

i.e., it is not affected by preference parameters.

Proof. The property follows from linearity of consumption policy functions in initial wealth which we first establish. We again simplify notation and drop the i and t indices. Recursive substitution from $j = J, \ldots, 1$, using that $\tilde{u}_J = \tilde{c}_J$ gives

$$\tilde{u}_1 = \left[\sum_{j=1}^{J} \tilde{\beta}^{j-1} \tilde{c}_j^{\frac{1-\theta}{\gamma}} \right]^{\frac{\gamma}{1-\theta}},$$

where $\tilde{\beta} = \beta(1+\lambda)^{\frac{1-\theta}{\gamma}}$. As for the resource constraint, write

$$\sum_{j=1}^{J} \tilde{y}_j \left(\frac{1}{1+r} \right)^{j-1} - \sum_{j=1}^{J} \tilde{c}_j \left(\frac{1}{1+r} \right)^{j-1} \geq 0,$$

where, in slight abuse of notation, we use $\tilde{y}_j$ to denote labor income during the working period and retirement income thereafter (see main text).

The Lagrangian writes as

$$\mathcal{L} = \left[\sum_{j=1}^{J} \beta^{j-1} \tilde{c}_j^{\frac{1-\theta}{\gamma}} \right]^{\frac{\gamma}{1-\theta}} + \lambda \left(\tilde{a}_1 + \sum_{j=1}^{J} \tilde{y}_j \left(\frac{1}{1+r} \right)^{j-1} - \sum_{j=1}^{J} \tilde{c}_j \left(\frac{1}{1+r} \right)^{j-1} \right).$$

First-order conditions give:

$$\tilde{\beta}^{j-1} \frac{1-\theta}{\gamma} \tilde{c}_j^{\frac{1-\theta-\gamma}{\gamma}} - \tilde{\lambda} \left(\frac{1}{1+r} \right)^{j-1} = 0$$

where $\tilde{\lambda} \equiv \lambda \left(\frac{\gamma}{1-\theta} \left[\sum_{j=1}^{J} \tilde{\beta}^{j-1} \tilde{c}_j^{\frac{1-\theta}{\gamma}} \right]^{\frac{\gamma}{1-\theta}-1} \right)^{-1}$. Using the FOC for any two ages j and $j+1$ gives the standard Euler equation

$$\frac{\tilde{c}_{j+1}}{\tilde{c}_j} = \left(\tilde{\beta}(1+r) \right)^{\frac{\gamma}{\theta+\gamma-1}} = \left(\tilde{\beta}(1+r) \right)^{\psi}.$$

where ψ is the intertemporal elasticity of substitution. We consequently have

$$\frac{\tilde{c}_j}{\tilde{c}_1} = (\beta(1+r))^{\psi(j-1)}.$$

Using this in the resource constraint, which holds with equality in the optimum, and defining by $\widetilde{pvi}_1$ total (human) wealth of the household, we get

$$\begin{aligned} \widetilde{pvi}_1 &\equiv \sum_{j=1}^{J} \tilde{y}_j \left(\frac{1}{1+r}\right)^{j-1} = \tilde{c}_1 \sum_{j=1}^{J} \frac{\tilde{c}_j}{\tilde{c}_1} \left(\frac{1}{1+r}\right)^{j-1} \\ \Leftrightarrow \qquad \widetilde{pvi}_1 &= \tilde{c}_1 \sum_{j=1}^{J} \left(\left(\tilde{\beta}(1+r)\right)^{\psi} \left(\frac{1}{1+r}\right)\right)^{j-1} = \tilde{c}_1 \sum_{j=1}^{J} b^{j-1} = \frac{1}{m_1}\tilde{c}_1 \end{aligned}$$

where $b \equiv \left(\tilde{\beta}(1+r)\right)^{\psi} \left(\frac{1}{1+r}\right)$ and $m_1 \equiv \left(\sum_{j=1}^{J} b^{j-1}\right)^{-1}$ is the marginal propensity to consume out of initial wealth in period 1. We accordingly get, for any age j, that

$$\tilde{c}_j = m_j \widetilde{pvi}_1, \text{ where } m_j \equiv \tilde{\beta}(1+r)^{\psi(j-1)} m_1.$$

Using this in the utility function we get

$$\tilde{u}_1 = \left[\sum_{j=1}^{J} \beta^{j-1} \left(m_j \widetilde{pvi}_1\right)^{\frac{1-\theta}{\gamma}}\right]^{\frac{\gamma}{1-\theta}} = \left[\sum_{j=1}^{J} \beta^{j-1} \left(m_j\right)^{\frac{1-\theta}{\gamma}}\right]^{\frac{\gamma}{1-\theta}} \widetilde{pvi}_1,$$

establishing linearity of the utility function in initial wealth.

Consequently, the CEV in partial equilibrium—where m_j does not change between any two policies A and B because it is only a function of the constant parameters r, β, ψ—is equal to the percentage change in wealth and given by

$$g_c^{PE} = \frac{\tilde{u}_1^A}{\tilde{u}_1^B} - 1 = \frac{\widetilde{pvi}_1^A}{\widetilde{pvi}_1^B} - 1.$$

□

B.5 Additional Proofs

Derivation of Equation (7). Here, we derive the stock and bond return in the quantitative model. Recall that $\overline{\kappa}_f$ is the exogenous and constant debt-equity ratio. First, we restate the relevant equation from Section 3.3,

$$K(z^t) = S(z^t) + B(z^t) = S(z^t)(1 + \overline{\kappa}_f), \tag{26}$$

where S and B denote the quantities of stock and bond, respectively. The return on capital then satisfies

$$r(z^t)K(z^t) = r(z^t)S(z^t)(1 + \overline{\kappa}_f).$$

The return on capital equals the standard first-order condition of the firm, as shown in equation (6b). Out of this total return on capital, bondholders receive

$$r_b(z^{t-1})B(z^t) = r_b(z^{t-1})\overline{\kappa}_f S(z^t),$$

where the bond return is determined one period ahead, since it is one-period risk-free. Stock holders receive the remainder, which is

$$r_s(z^t)S(z^t) = r(z^t)S(z^t)(1 + \overline{\kappa}_f) - r_b(z^{t-1})\overline{\kappa}_f S(z^t).$$

From the last equation, we immediately get (7). □

Proof of Equation (8). The property follows from homotheticity of Epstein-Zin preferences. To prove it we proceed by induction. We look at two alternative (expected) consumption streams $\tilde{c}^A$ and $\tilde{c}^B$. One can think of them as optimal consumption under policy regime A and B. We ask how big the percentage increase of consumption stream $\tilde{c}^A$ in each period has to be to reach the same utility level as reached for consumption stream $\tilde{\tilde{c}}^B$. For sake of simplicity we drop indices t and i and adopt the notation $\tilde{u}_j^X \equiv \tilde{u}_j(c^X)$ for $X \in \{A, B\}$.

1. *Induction claim:* At each age j we have that

$$\tilde{u}_j^B = (1 + g_c)\tilde{u}_j^A.$$

2. *Induction start*: For our Epstein-Zin utility specification (cf. Section 3.2), at age J we have that

$$\tilde{u}_J^A = \tilde{c}_J^A \text{ and } \tilde{u}_J^B = \tilde{c}_J^B.$$

   Hence, by the induction claim, we get

$$\tilde{u}_J^B = (1 + g_c)\tilde{u}_J^A = (1 + g_c)\tilde{c}_J^A$$

and, correspondingly,

$$\begin{aligned}\tilde{u}^B_{J-1} &= \left[(\tilde{c}^B_{J-1})^{\frac{1-\theta}{\gamma}} + \beta\left(\mathbb{E}_{J-1}(\tilde{u}^B_J)^{1-\theta}\right)^{\frac{1}{\gamma}}\right]^{\frac{\gamma}{1-\theta}} \\ &= (1+g_c)\tilde{u}^A_{J-1}.\end{aligned}$$

3. *Induction step*: Using the induction claim for any period $j < J-1$ we therefore have:

$$\begin{aligned}\tilde{u}^B_j &= \left[(\tilde{c}^B_j)^{\frac{1-\theta}{\gamma}} + \beta\left(\mathbb{E}_j(\tilde{u}^B_{j+1})^{1-\theta}\right)^{\frac{1}{\gamma}}\right]^{\frac{\gamma}{1-\theta}} \\ &= (1+g_c)\tilde{u}^A_j.\end{aligned}$$

□

Derivation of Equation (11). Denote by $\gamma_c = \frac{E\left[\tilde{C}^B|\tau=2\%\right]}{E\left[\tilde{C}^A|\tau=0\%\right]} = 1 + \Delta_c$ the consumption growth factor, where Δ^c is the consumption growth rate. Take the definition of g_c and divide all individual consumption allocations, $\tilde{c}^B$, by the consumption growth factor so that mean consumption is the same in both economies. This gives the welfare benefits from changes of the distribution as

$$g^{distr}_c = \frac{E\left[\tilde{v}_1\left(\frac{1}{\gamma_c}\tilde{c}^B\right)|\tau=2\%\right]}{E\left[\tilde{v}_1(\tilde{c}^A)|\tau=0\%\right]} - 1 = \frac{1}{\gamma_c}\frac{E\left[\tilde{v}_1\left(\tilde{c}^B\right)|\tau=2\%\right]}{E\left[\tilde{v}_1(\tilde{c}^A)|\tau=0\%\right]} - 1 = \frac{1}{\gamma_c}(1+g_c) - 1,$$

where the second equality follows from homotheticity. The g^{mean}_c is then given by

$$g^{mean}_c = g_c - g^{distr}_c = \frac{1+g_c}{\gamma_c}\Delta_c.$$

□

C Supplementary Computational Appendix

C.1 Overview

The numerical solution follows Krusell and Smith (1997, 1998), Storesletten, Telmer, and Yaron (2007) and Gomes and Michaelides (2008). The algorithm consists of the following steps, details of which are given in the next subsections.

1. Choose arguments and a functional form for the approximate law of motion, and make an initial guess for its coefficients.

2. Given the approximate law of motion, solve the household's problem.
3. Simulate the economy using the obtained optimal policy functions. In every period, compute the market clearing prices.
4. Update the coefficients of the approximate law of motion by running a regression on the simulated aggregate statistics.
5. If the coefficients have converged, and the R^2 of the regression is sufficiently high, stop, else go to 2.
6. Repeat steps 1 to 5 for different arguments and functional forms of the law of motion. Select the one with the highest R^2.
7. Given the functional form for the approximate law of motion that achieved the best fit, calibrate the economy to match the targets.
    (a) Provide an initial guess for the parameters to be calibrated.
    (b) Given the parameters, repeat steps 2 to 5.
    (c) Calculate the target statistics from the simulations. If they are close to the targets in the data, stop, else update the guess for the parameters and go to 7b.
8. Given the calibrated parameters, increase the social security contribution rate and compute the new general equilibrium by repeating steps 2 to 5.
9. Compute the welfare gains of the experiment in general equilibrium from the simulated variables of the first and the second economy.
10. Given the approximate laws of motion and the simulated prices of the first economy, perform the risk decomposition analysis.
    (a) Given the approximate law of motion of the first economy, solve the household's problem.
    (b) Given the simulated prices of the first economy, simulate the economy using the obtained optimal policy functions. (Do not compute market clearing prices.)
    (c) Increase the social security contribution rate and repeat steps 10a and 10b.
    (d) Compute the welfare gains of the experiment in partial equilibrium (PE) from the simulated variables of the pre-experiment PE and the post-experiment PE.

(e) If this was the no-risk, deterministic economy, stop, else turn off a risk and repeat steps 10a to 10d.

The numerical solution is implemented in Fortran and parallelized, running on 24 cores.

C.2 Solving for the approximate law of motion

The idea behind the Krusell-Smith-method (1997, 1998) is to approximate the infinite dimensional distribution, Φ, by a finite number of statistics. The household then uses a law of motion of these statistics, $\hat{H}(\cdot)$, as an approximation to the true law of motion of the distribution, $H(\Phi, z, z')$. The statistics have to enable the household to forecast the prices that it needs to solve its optimization problem. We follow Krusell and Smith (1997), Storesletten, Telmer, and Yaron (2007), and Gomes and Michaelides (2008) and choose mean aggregate capital, k, together with a second variable to forecast the bond return. As this second variable, we choose the expected equity premium, $\mu = \mathbb{E}\left(r'_s - r'_b\right)$, cf. Storesletten, Telmer, and Yaron (2007).[49] Thus, the approximate law of motion becomes

$$\{k'(z'), \mu'(z')\} = \hat{H}(k, \mu, z, z').$$

The functional form for $\hat{H}$ that gives the best approximation in our baseline economy is

$$\ln k_{t+1} = \psi^k_{0,z} + \psi^k_{1,z} \ln k_t + \psi^k_{2,z} (\ln k_t)^2, \tag{27a}$$

$$\mu_{t+1,z'} = \psi^\mu_{0,z'} + \psi^\mu_{1,z'} \ln k_{t+1} + \psi^\mu_{2,z'} (\ln k_{t+1})^2, \tag{27b}$$

which is similar to the best fit regression found by Storesletten, Telmer, and Yaron (2007). Note that the forecast of capital, $\ln k_{t+1}$, enters as a regressor in eq. (27b). Effectively, the forecast for $\mu_{t+1,z'}$, which is conditional on z', depends on $\ln k_t$ and z through the forecast of $\ln k_{t+1}$. The discrete, aggregate shock, z, can take four values, so that we estimate eight equations. Therefore, we report eight coefficients of determination, which for the baseline economy are $R^2_k = \{0.9998, 0.9999, 0.9998, 0.9998\}$ and $R^2_\mu = \{0.9918, 0.9945, 0.9584, 0.9695\}$. For the other economies, the R^2s are always higher.[50]

[49] We choose μ instead of the bond price because this enables us to avoid $E(r_b) > E(r_s)$ by construction. This is desirable because such a situation would never arise in equilibrium.

[50] For example, for the equity premium calibration with $IES = 0.5$, the coefficients of determination are $R^2_k = \{0.9999, 0.9999, 0.9999, 0.9999\}$ and $R^2_\mu = \{0.9961, 0.9968, 0.9944, 0.9949\}$. This economy is the closest to Storesletten, Telmer, and Yaron (2007) and Gomes and Michaelides (2008), and the R^2 are very close to the ones reported there.

To find the coefficients, we solve $g(\Psi) = \Psi - \tilde{\Psi}(\Psi)$, where Ψ collects all the coefficients, i.e. $\Psi = \left\{\psi_{l,z}^{m}\right\}_{l=\{0,1,2\},z=\{1,2,3,4\},m=\{k,\mu\}}$. To solve this nonlinear equation system, a multidimensional Broyden algorithm is used. During the solution, we normalize (and subsequently de-normalize) the coefficients around unity. For these coefficients around unity, the convergence criterion is $\max\{|g(\Psi)|\} < 1.0^{-7}$. The Newton-like update steps are limited to a small length, and backtracking is used to find an update, if the first step was too large.[51]

C.3 Solving the household's problem

First, we rewrite the household problem in terms of cash-on-hand, $\tilde{x}$. This reduces the state space by one dimension, so that the idiosyncratic state consists of $(\tilde{x}, e)$. Second, we recast the two control variables bond, $\tilde{b}'$, and stock, $\tilde{s}'$, as total savings, $\tilde{a}'$, and the portfolio share invested in stock, κ. This enables us to employ the endogenous grid method proposed by Carroll (2006), as detailed below. And third, we replace the distribution, Φ, by the approximation discussed in the previous section, so that the aggregate state consists of (k, μ, z). With a slight abuse of notation,[52] the optimization problem in recursive form then reads

$$
\begin{aligned}
&\tilde{v}_j(\tilde{x}, e; k, \mu, z) \\
&= \max_{\tilde{c}>0, \tilde{a}', \kappa} \begin{cases} \left(\tilde{c}^{\frac{1-\theta}{\gamma}} + \tilde{\beta}\left(\sum_{z'}\sum_{e'} \pi_z(z'|z)\pi_e(e'|e)\, \tilde{v}_{j+1}^{1-\theta}\left(\tilde{x}', e'; \hat{H}(k, \mu, z, z'), z'\right)\right)^{\frac{1}{\gamma}}\right)^{\frac{\gamma}{1-\theta}} \\ \tilde{c} & \text{if } j = J \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&\text{s. t.} \\
&\tilde{x}' = \tilde{a}' \frac{(1 + {r_b}' + \kappa(r_s' - r_b'))}{1 + \lambda} + \tilde{y}_{j+1}' \;, \\
&\tilde{a}' \geq 0 \qquad \text{if } \; j = J \;,
\end{aligned}
$$

[51]The Newton-like update step is $\Psi_{i+1} = \Psi_i - sJ(\Psi)^{-1}g(\Psi)$, where $J(\Psi)$ is a finite-difference approximation to the Jacobi matrix of the system of equations and s determines the maximum step length.

[52]Technically, some variables would need to be renamed, e.g. $\tilde{y}$ to $\tilde{\tilde{y}}$, because the state space is now different than the one in Definition 4. For sake of readability, we do not change the notation.

where $\tilde{\beta} = \beta(1+\lambda)^{\frac{1-\theta}{\gamma}}$, $r_s' = r_s(\hat{H}(k,\mu,z,z'),z')$, $r_b' = r_b(\hat{H}(k,\mu,z,z'))$, and income in the next period is given by

$$\begin{aligned}\tilde{y}_{j+1}' &= \tilde{y}_{j+1}(e',\hat{H}(k,\mu,z,z'),z') \\ &= \begin{cases}(1-\tau)\tilde{w}(\hat{H}(k,\mu,z,z'),z')\epsilon_{j+1}\eta(e',z') & \text{if } j+1<j_r, \\ \tilde{y}_{ss}(\hat{H}(k,\mu,z,z'),z') & \text{else.}\end{cases}\end{aligned}$$

The budget constraint contains a growth adjustment of $\frac{1}{1+\lambda}$, because x' is cash on hand at the beginning of next period, while a' is the savings choice made in the previous period. In contrast, the budget constraint in the equilibrium definition of Section 3.5 contains only contemporaneous variables, i.e., states and choices in the current period, so that no growth adjustment is needed there.

Applying the envelope theorem and simplifying we get the two first-order-conditions[53]

$$\mathbb{E}\left[\tilde{v}_{j+1}(\cdot)^{\frac{(1-\theta)(\gamma-1)}{\gamma}}(\tilde{c}')^{\frac{1-\theta-\gamma}{\gamma}}(r_s'-r_b')\right] = 0\ , \tag{28a}$$

$$\tilde{c}=\left(\tilde{\beta}\frac{1+r_b{}'}{1+\lambda}\left(\mathbb{E}\left[\tilde{v}_{j+1}(\cdot)^{1-\theta}\right]\right)^{\frac{1-\gamma}{\gamma}}\mathbb{E}\left[\tilde{v}_{j+1}(\cdot)^{\frac{(1-\theta)(\gamma-1)}{\gamma}}(\tilde{c}')^{\frac{1-\theta-\gamma}{\gamma}}\right]\right)^{\frac{\gamma}{1-\theta-\gamma}}. \tag{28b}$$

To solve for the optimal choices $(\tilde{c},\tilde{a}',\kappa)$, we apply a variant of the endogenous grid method proposed first by Carroll (2006). In fact, essentially we follow a simplified version of the two-step procedure of Hintermaier and Koeniger (2010). The exogenous grid is defined on total assets in the next period, $\tilde{a}'$. For a given grid-point $\tilde{a}_i'$, we first solve eq. (28a) for the portfolio share κ using Brent's root-finding method. Then, given $\tilde{a}_i'$ and the corresponding $\kappa(\tilde{a}_i')$, we use eq. (28b) to get the optimal consumption, $\tilde{c}_i(\tilde{a}_i')$. Finally, the budget constraint $\tilde{x} = \tilde{c} + \tilde{a}'$ gives us the endogenous grid-point $\tilde{x}_i$ that corresponds to the optimal choices $(\tilde{a}_i',\tilde{c}_i)$.

When evaluating the expectations, we interpolate $\tilde{v}_{j+1}$ and $\tilde{c}'$ by multidimensional linear interpolation in the continuous states $\tilde{x},k,\mu$. The aggregate shock z and the idiosyncratic shock $\eta(e,z)$ are both discrete and follow a discrete Markov chain. As discussed in Section 4.2, our specification of the stochastic labor income process has a persistent component, ν ,and a transitory shock, ε, so that we have $\eta(e,z) = \nu(e,z)\varepsilon(e)$. We construct the Markov transition matrix of $\nu(e,z)$ with the Rouwenhorst method (Kopecky and Suen (2010)), which makes it straightforward to implement the countercyclical cross-sectional variance, CCV, because the variances affect only the grid and not the transition matrix, which in turn is determined purely by ρ. We discretize the transitory shocks ε using Gauss-Hermite quadrature.

[53]See Weil (1989) for the envelope theorem with recursive Epstein-Zin preferences.

As is standard in life-cycle models, we iterate backwards, starting with the last generation J, for which the solution is $\tilde{c}_J = \tilde{x}_J$, since they do not leave bequests. In the backwards iteration, we construct age-dependent, exogenous grids $\left\{\tilde{a}'_{i,j}\right\}_{i,j}$ to improve the approximation quality. The solution is parallelized in the dimension k, so that for each generation, the solution for all values of k is computed in parallel.

We discretize the state space in the following way. The continuous state variables cash-on-hand, $\tilde{x}$, aggregate capital, k, and equity premium, μ, have 20, 18, and 10 grid-points, respectively. The discrete state variables, which are the number of generations, J, the idiosyncratic shock, e, and the aggregate shock, z, have 58, 4, and 4 grid-points, respectively, and we use 4 points for the Gauss-Hermite quadrature (for the transitory shocks). We check that this is sufficient by doubling each of the grid-points in turn and find no change to our results. The first-order-condition in eq. (28a) is solved to an accuracy of 1.0^{-10}.

C.4 Simulating the economy

We simulate the economy 24 times for 4000 periods each time and throw away the first 1000 periods, so that we are left with a total 72.000 simulation periods.[54] In each period, we record the aggregates, the life-cycle profiles, and the distribution. The aggregates are needed to estimate the laws of motion, and to calibrate the economy. Like in the solution of the household problem, the optimal policy functions are interpolated in the dimensions of the aggregate states k, μ by multidimensional linear interpolation.

The distribution over households is normalized to a mass of one. We do not simulate many, discrete household units; instead we keep the continuum of households and approximate the distribution with a histogram as proposed by Young (2010). As described in Section 3.1, the Law of Large Numbers implies that $\pi_e(e'|e)$ represents the fraction of the population moving from idiosyncratic state e to e'. Therefore, we get a nearly exact approximation in that dimension. In the cash-on-hand dimension, the distribution is discretized on a much finer grid than the policy functions obtained in the household solution, as proposed by Ríos-Rull (1999). This finer discretization improves the approximation quality substantially and helps in ensuring that no households are stuck on the bounds of the distribution. If the lowest or the highest points of the distribution have positive mass, then the cash-on-hand grid is extended and the discretization is made finer.

[54] We found that a large number of simulation periods is necessary for the distribution to converge in the sense that increasing the number of simulation periods does not change the results. In particular, we found that for less than 30.000 simulation periods, the means and standard deviations of the aggregates as well as the estimates of the laws of motion are still sensitive to the number of periods.

In each period, the beginning-of-period distribution is iterated forward by using the computed optimal policy functions and the realizations of the shock. For a given cash-on-hand at the beginning of the period, the implied cash-on-hand in the following period will almost always lie between two grid points. Since we are dealing with a continuum of households, we assign a fraction $(1 - f)$ to the lower grid point and f to the upper grid point of the interval which contains the implied cash-on-hand, where f is the distance to the lower grid point.[55]

In each period t, we calculate the market-clearing prices.[56] The current stock return, $r_s(\Phi_t, z_t)$ is given by the contemporaneous aggregate capital and aggregate shock. The current bond return, $r_b(\Phi_t)$, is determined one period before by the bond market clearing condition. We compute it with a nonlinear equation solver to an accuracy of 1.0^{-8}.

We make sure that the grid for the aggregate states is large enough by checking whether the realized values lie on the bounds of the grid. If they do, the grid is increased. To get good initial guesses for the bounds of the aggregate grids and the distribution over households, we compute a degenerate equilibrium, where the realization of the aggregate shocks in the simulations is always equal to their mean. We call this a mean-shock equilibrium.

To check the accuracy of the solution, we compute in each period the 'aggregation error' and the 'income error'. The aggregation error $e_t^{agg} = \frac{Y_t - C_t - I_t}{Y_t}$ says by how much the aggregate budget constraint is violated due to interpolation and aggregation errors, expressed in percent of output. For all economies, the maximum aggregation error is in the order of 1.0^{-6} and the average is in the order of 2.0^{-9}. The income error comes from Euler's formula, which says that total output must equal total factor income. Again expressed in terms of output, we find that it never exceeds 1.0^{-14}.

C.5 Calibrating the economy

The calibration procedure is cast as a system of nonlinear equations. Let $\mathcal{T}$ denote the target statistics in the data and $\mathcal{P}$ the model parameters to be calibrated. For given $\mathcal{P}$, $\hat{\mathcal{T}}(\mathcal{P})$ are the model-generated statistics, which we get from the simulations. Then the calibration procedure tries to find a root of $\mathcal{T} - \hat{\mathcal{T}}(\mathcal{P}) = 0$. We use Broyden's multidimensional secant method to solve the system to an accuracy of 1.0^{-4}.

[55] For details, see Young (2010).

[56] Algan, Allais, Den Haan, and Rendahl (2014) stress the importance of ensuring market clearing during the simulations.

C.6 Evaluating the conditions for dynamic efficiency

After having calibrated the economy, we evaluate the sufficient conditions of Definition 3 in each period of our simulations (after discarding burn-in). Conditional on having reached a high bond return in a date-event z^t, $r_b(\Phi_t) > (1+n)(1+g)$, we compute market clearing prices for all possible aggregate shocks next period, $z_{t+1} \in \mathcal{Z}$.[57] We then check whether there exist two states $\tilde{z}_{t+1}, \tilde{\tilde{z}}_{t+1}$, such that (i) the economy remains in a high bond return equilibrium in both corresponding date-events next period and (ii) the stock return fluctuates enough in these states relative to this period's high bond return. If this is the case condition (a) is fulfilled.

As for condition (b), we define a counter which is initialized to zero and increased by one every period. If the economy reaches a high bond return equilibrium, the value of the counter is saved, and we start counting again from zero. If the maximum count is smaller than the number of simulated periods, then condition (b) is fulfilled. To provide more information we report in Tables 2 and 12 the maximum and the average count.

D Supplementary Calibration Appendix

D.1 Households

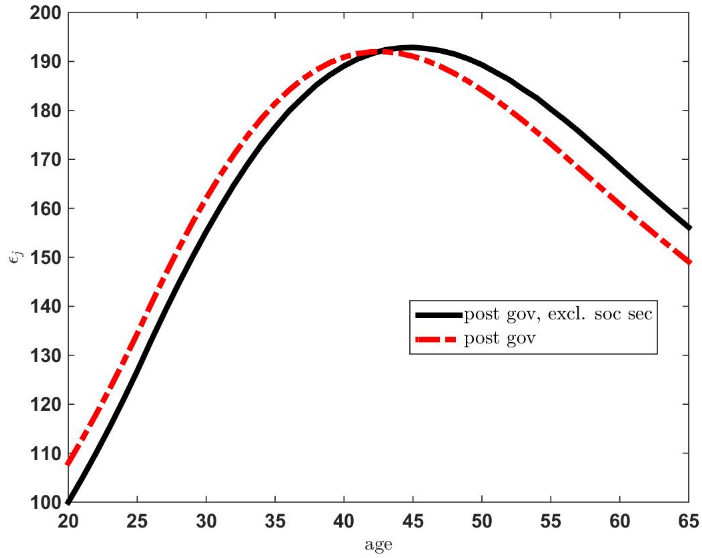
We base our estimates of the earnings process on Busch and Ludwig (2017) using PSID data of household post government earnings excluding contributions to social security. To get consistent earnings measures the sample is restricted to years from 1977 to 2012. Household pre-government earnings is the sum of labor earnings of households' head and spouses augmented by 50% of payroll taxes excluding social security. Post-government earnings is derived from this measure by adding transfers and deducting taxes (calculated from TAXSIM), again excluding contributions to social security. A second earnings measure used for our scenario with social security ($BL_{\tau=9.5\%}$) in Section 5.3 also takes contributions to social security into account.

Busch and Ludwig (2017) adopt the standard strategy to first decompose a household's log earnings into a deterministic and a stochastic component. The estimates of the age specific productivity profile ϵ_j are taken from the deterministic component. Figure 3 displays the profile for both our earnings measures ("post gov, excl. soc sec" and "post gov").

To estimate the stochastic earnings process, Busch and Ludwig (2017) first classify contraction years on the basis of NBER recession indicators, which, due to the sluggish adjustment of the labor market, is expanded by years of upward trending unemployment rates as in Guvenen,

[57] Since the Markov transition matrix π_z has non-zero entries everywhere, all z^{t+1} can be reached from z^t with positive probability.

Figure 3: Life-cycle Productivity



Notes: Age-specific productivity profile ϵ_j. *Notes:* Own estimates based on Busch and Ludwig (2017).

Ozkan, and Song (2014). All years that are not accordingly classified as a recession are classified as a boom. Using their procedure and the income measures described above, our earnings process (12) is estimated with results reported in Section 4.2.[58]

D.2 Firms

To estimate α, we take data on total compensation of employees (NIPA Table 1.12) and deflate it with the GDP deflator (NIPA Table 1.1.4). In the numerator, we adjust GDP (NIPA Table 1.1.5), again deflated by the GDP deflator, by nonfarm proprietors' income and other factors that should not be directly related to wage income. Without these adjustments, our estimate of α would be considerably higher (at $\alpha = 0.43$).

To measure capital, we take the stock of fixed assets (NIPA Table 1.1), appropriately deflated. We relate this to total GDP.

We determine the growth rate of technology λ by estimating the Solow residual from the production function, given our estimate of α, our measure for capital, and a measure of labor supply determined by multiplying all full- and part-time employees in domestic employment (NIPA Table 6.4A) with an index for aggregate hours (NIPA Table 6.4A). Notice that we ignore

[58]Again, we thank Christopher Busch for providing us with the estimates.

age-specific productivity which should augment our measure of employment. We then fit a linear trend specification to the Solow residual. Acknowledging the labor augmenting technological progress specification, this gives our point estimate.

D.3 Aggregate Risk

We first provide details on how we construct the transition matrix and the values for the aggregate technology and depreciation shocks, $(\zeta(z), \delta(z))$. Both $\zeta(z)$ and $\delta(z)$ can each take a high or a low value. We let

$$\zeta(z) = \begin{cases} 1 - \bar{\zeta} \text{ for } z \in z_1, z_2 \\ 1 + \bar{\zeta} \text{ for } z \in z_3, z_4 \end{cases} \quad \text{and } \delta(z) = \begin{cases} \delta_0 + \bar{\delta} \text{ for } z \in z_1, z_3 \\ \delta_0 - \bar{\delta} \text{ for } z \in z_2, z_4. \end{cases} \tag{29}$$

Set up in this way, z_1 corresponds to a low wage and a low return, while z_4 corresponds to a high wage and a high return. We speak of $z \in z_1, z_2$ as a *recessions* in the sense that these represent states in which aggregate wage shocks are low.

To calibrate the entries in the transition matrix, denote the transition probability of remaining in the low technology state by $\pi^\zeta = \pi(\zeta' = 1 - \bar{\zeta} \mid \zeta = 1 - \bar{\zeta})$. Assuming that the transition of technology shocks is symmetric, we then have $\pi(\zeta' = 1 + \bar{\zeta} \mid \zeta = 1 + \bar{\zeta}) = \pi^\zeta$ and $1 - \pi^\zeta = \pi(\zeta' = 1 - \bar{\zeta} \mid \zeta = 1 + \bar{\zeta}) = \pi(\zeta' = 1 + \bar{\zeta} \mid \zeta = 1 - \bar{\zeta})$.

To govern the correlation between technology and depreciation shocks, let the probability of being in the high (low) depreciation state conditional on being in the low (high) technology state be $\pi^\delta = \pi(\delta' = \delta_0 + \bar{\delta} \mid \zeta' = 1 - \bar{\zeta}) = \pi(\delta' = \delta_0 - \bar{\delta} \mid \zeta' = 1 + \bar{\zeta})$, where the second equality follows from assuming symmetry of the matrix. We then have that the transition matrix of aggregate states follows from the corresponding assignment of states in (29) as

$$\pi_z = \begin{bmatrix} \pi^\zeta \cdot \pi^\delta & \pi^\zeta \cdot (1 - \pi^\delta) & (1 - \pi^\zeta) \cdot (1 - \pi^\delta) & (1 - \pi^\zeta) \cdot \pi^\delta \\ \pi^\zeta \cdot \pi^\delta & \pi^\zeta \cdot (1 - \pi^\delta) & (1 - \pi^\zeta) \cdot (1 - \pi^\delta) & (1 - \pi^\zeta) \cdot \pi^\delta \\ (1 - \pi^\zeta) \cdot \pi^\delta & (1 - \pi^\zeta) \cdot (1 - \pi^\delta) & \pi^\zeta \cdot (1 - \pi^\delta) & \pi^\zeta \cdot \pi^\delta \\ (1 - \pi^\zeta) \cdot \pi^\delta & (1 - \pi^\zeta) \cdot (1 - \pi^\delta) & \pi^\zeta \cdot (1 - \pi^\delta) & \pi^\zeta \cdot \pi^\delta \end{bmatrix}.$$

Now we discuss the empirical correlation of TFP and stock returns, $\sigma(\zeta_t, r_{s,t})$, a second-stage calibration target. Linear detrending of the data, as done, e.g., by Krueger and Kubler (2006), results in $\sigma(\zeta_t, r_{s,t}) < 0$ as well as a negative correlation of wages and asset returns, i.e., $\sigma(w_t, r_{s,t}) < 0$. Not only does this seem counter to economic intuition in an annual RBC model, but our estimate for $\sigma(\zeta_t, r_{s,t})$ is also statistically insignificant. Assuming instead a unit root process for the log of TFP and detrending by first differences yields a highly significant positive correlation of $\sigma(\zeta_t, r_{s,t}) = 0.50$ (p-value 0.00).[59] Now also $\sigma(w_t, r_{s,t})$ is positive and significant with $\sigma(w_t, r_{s,t}) = 0.306$ (p-value 0.025), which coincides with our economic intuition as we would expect these variables to co-move over the cycle. Our model, however, features a linear trend, not a unit root. We therefore translate these moments to be consistent with a deterministic trend specification.

D.4 Calibration of Single Risk Economies

Table 10 summarizes the second-stage parameters, i.e., the parameters that are jointly calibrated. The remaining first-stage parameters take the same value as in the baseline, see Table 1. Table 10 also displays the targeted moments for these economies. For comparison, the table includes the corresponding values of the baseline (BL).

Table 10: The Role of Both Risks: Parameters and Moments

Parameter	θ	β	δ_0	$\bar{\delta}$	π^{δ}
BL	3.00	0.987	0.102	0.080	0.887
AR-$only$	15.10	0.994	0.076	0.111	0.833
IR-$only$	3.00	0.974	0.079	0.000	NA
No-$risk$	3.00	0.998	0.079	0.000	NA

Moment	ς	μ	$E\left[\frac{K}{Y}\right]$	$E\left[r_b\right]$	$\sigma\left(\frac{\Delta C_{t+1}}{C_t}\right)$	$\sigma(r_s)$	$\sigma(\zeta, r_s)$
BL	0.076	0.008	2.65	0.023	0.030	0.107	0.500
AR-$only$	0.351	0.056	2.65	0.023	0.040	0.168	0.500
IR-$only$	0.000	0.000	2.65	0.042	NA	NA	NA
No-$risk$	0.000	0.000	2.65	0.042	NA	NA	NA

Notes: BL: baseline calibration with $\theta = 3$; AR-$only$: economy with only aggregate risk, calibrated to match equity premium; IR-$only$: economy with only idiosyncratic risk; No-$risk$: deterministic economy. $\varsigma = \frac{E[r_{s,t}-r_{b,t}]}{\sigma[r_{s,t}-r_{b,t}]}$: Sharpe ratio; $\mu = E\left[r_{s,t} - r_{b,t}\right]$: equity premium; $E\left[\frac{K}{Y}\right]$: average capital-output ratio; $E\left[r_b\right]$: average bond return; $\sigma\left(\frac{\Delta C_{t+1}}{C_t}\right)$: standard deviation of aggregate consumption; $\sigma(r_s)$: standard deviation of stock returns; $\sigma(\zeta, r_s)$: correlation of TFP shocks and stock returns.

[59] Observe that calibrating the model to match this moment explicitly is more conservative with regard to the return implications of recessions than the assumption of Storesletten, Telmer, and Yaron (2007) who assume a perfect negative correlation of TFP and depreciation shocks.

D.5 Calibration for Sensitivity Analyses

The calibrated parameters and targeted moments for the various scenarios we consider are summarized in Table 11. For comparison, the table includes the corresponding values of the baseline (BL).

For the calibration of the earnings process in the $\tau = 9.5\%$ economy, we take the Busch and Ludwig (2017) estimates of an income process cum social security. The age productivity profile is shown in Figure 3. For the stochastic part of the income process the estimates are $\rho = 0.966$, and a conditional variance, $\sigma_{\nu}^2(z_t)$, of 0.024 in recessions and 0.01 in booms, and $\sigma_{\varepsilon}^2 = 0.099$. Setting $\tau = 9.5\%$ in this economy gives an average net pension benefit level $\frac{Ey^{ss}(z^t)}{(1-\tau)Ew(z^t)}$, of 49.8%.

E Supplementary Results Appendix

E.1 Dynamic Efficiency

Table 12 reports the results for the sufficient conditions according to Definition 3 for all model variants with aggregate risk considered in the main text, always for the scenario without social security ($\tau = 0\%$), respectively with a contribution rate of $\tau = 7.5\%$. The number of simulated periods used is $72\,000$. Throughout, we conclude that all considered economies are dynamically efficient. The lowest confidence for this finding of at least 90.0% applies to the economy $BL_{IES=1.5}$.

E.2 On the Importance of Modeling both Risks: Results Without Re-Calibration

Complementing our analysis in Section 5.2 we summarize in Table 13 results for economies with only one form of risk, respectively for the deterministic economy, when we do not recalibrate. As in Table 8, there are welfare losses in the AR-*only* economy.

The IR-*only* economy is dynamically inefficient. The risk-free return is at 2.86%, compared to the implicit return of social security of 2.9%. That is why we find a small welfare gain in general equilibrium. Despite dynamic inefficiency, there are welfare losses from crowding out, cf. our discussion in Section 3.8 on the relationship between the costs from crowding out and dynamic efficiency in heterogeneous agent economies. Finally, the deterministic economy has a risk-free interest rate of 5.56% and we therefore continue to find welfare losses from the introduction of social security.

Table 11: Sensitivity Analysis: Parameters and Moments

Parameter	θ	β	δ_0	$\bar{\delta}$	π^δ
$IES = 0.5$					
BL	3.00	0.987	0.102	0.080	0.887
SR	11.10	0.987	0.020	0.045	0.829
EP	5.51	0.987	0.000	0.114	0.830
$BL_{\tau=7.5\%}$	3.00	1.014	0.105	0.088	0.893
$IES = 1.5$					
$BL_{IES=1.5}$	3.00	0.977	0.099	0.039	0.888
$SR_{IES=1.5}$	12.20	0.977	0.019	0.038	0.829
$EP_{IES=1.5}$	5.60	0.977	0.001	0.115	0.835

Moment	ς	μ	$E\left[\frac{K}{Y}\right]$	$E\left[r_b\right]$	$\sigma\left(\frac{\Delta C_{t+1}}{C_t}\right)$	$\sigma(r_s)$	$\sigma(\zeta, r_s)$
$IES = 0.5$							
BL	0.076	0.008	2.65	0.023	0.030	0.107	0.500
SR	0.333	0.020	5.80	0.023	0.030	0.067	0.500
EP	0.357	0.056	7.67	0.023	0.066	0.168	0.500
$BL_{\tau=7.5\%}$	0.069	0.007	2.65	0.023	0.030	0.116	0.500
$IES = 1.5$							
$BL_{IES=1.5}$	0.051	0.003	2.65	0.023	0.030	0.052	0.500
$SR_{IES=1.5}$	0.333	0.018	5.90	0.023	0.030	0.057	0.500
$EP_{IES=1.5}$	0.356	0.056	6.57	0.023	0.089	0.168	0.500

Notes: BL: baseline calibration with $\theta = 3$; SR: scenario matching Sharpe ratio; EP: scenario matching equity premium. $\varsigma = \frac{E[r_{s,t}-r_{b,t}]}{\sigma[r_{s,t}-r_{b,t}]}$: Sharpe ratio; $\mu = E\left[r_{s,t} - r_{b,t}\right]$: equity premium; $E\left[\frac{K}{Y}\right]$: average capital-output ratio; $E\left[r_b\right]$: average bond return; $\sigma\left(\frac{\Delta C_{t+1}}{C_t}\right)$: standard deviation of aggregate consumption; $\sigma(r_s)$: standard deviation of stock returns; $\sigma(\zeta, r_s)$: correlation of TFP shocks and stock returns.

E.3 Other Sensitivity Analyses

Table 14 summarizes results for other sensitivity analyses. Throughout, we stick to the calibration strategy of the conservative baseline model—we accordingly denote all these scenarios with prefix BL—and vary selected parameters, respectively alter modeling choices. We accordingly recalibrate to match the calibration targets of the baseline, assuming an IES of 0.5.

First, we vary risk aversion in economies BL-$\theta = 2$ and BL-$\theta = 4$. As expected, increasing θ increases the overall welfare gains, increases the share of gains attributable to the insurance effect, decreases the share attributable to the mean effect and increases the share of

Table 12: Dynamic Efficiency in Baseline and Other Economies

	Condition (a)		Condition (b)	
	High Bond Returns	Conditional Violation	Max. Periods	Avg. Periods
$IES = 0.5$				
BL	38.1%	4.7%	120	11.5
SR	34.8%	6.0%	158	13.5
EP	45.1%	0.8%	134	10.8
$AR\text{-}only$	40.6%	3.9%	103	7.0
$BL_{\tau=7.5\%}$	41.2%	3.1%	120	10.9
$IES = 1.5$				
$BL_{IES=1.5}$	31.8%	10.0%	120	12.1
$SR_{IES=1.5}$	36.0%	6.6%	124	12.0
$EP_{IES=1.5}$	46.3%	0.8%	120	9.9

Notes: Test results for dynamic efficiency conditions, cf. Definition 3, before introducing, resp. expanding, the social security system (i.e., for economies with $\tau = 0\%$, resp. $\tau = 7.5\%$). High bond returns: fraction of high bond return equilibria in which $1 + r_b(z^t) > (1 + n)(1 + \lambda)$. Conditional violation: Violation of conditions (a)(i) and (a)(ii), conditional on being in a high bond equilibrium. Avg., resp. max., periods: average, resp. maximum, number of simulation periods to reach high bond equilibrium. The number of simulated and tested periods is 72 000 in each scenario.

the interactions; and vice versa for decreasing θ.

Next, we consider a number of experiments regarding different model elements. In scenario $BL\text{-}\sigma_\varepsilon^2 = 0$ we switch off the variance of transitory labor income shocks to show that our results are not driven by this element. Scenario $BL\text{-}\bar{\kappa} = 0$ sets the debt-equity-ratio to zero (so that firms are purely equity financed), thereby decreasing the equity premium from 0.76% in the baseline calibration to 0.46%, which slightly decreases the overall welfare gains from social security. Targeting a higher capital-output ratio of 3 in scenario $BL\text{-}\frac{K}{Y} = 3$ yields higher welfare benefits. The reason is that this requires a higher discount factor (of $\beta = 0.992$ rather than $\beta = 0.987$), which increases the welfare benefits from social security in line with the predictions from the extended simple model in Appendix B.2. Finally, we consider an experiment where we shut down the depreciation shocks in the conservative baseline calibration, $BL\text{-}\bar{\delta} = 0$. In this experiment we recalibrate the variance of TFP shocks (which in all other experiments is taken as a first-stage parameter) to match the consumption volatility. In this experiment, the partial equilibrium gains from insurance against aggregate risks decrease from 2% to 1.4%, which leads to a decrease of the gains from insurance against the interactions of risks and to a

Table 13: The Role of Both Risks without Re-Calibration

	Consumption equivalent variation, g_c			g_c^{distr}	g_c^{mean}
Scenario	GE	PE	CO	GE	
$AR\text{-}only$	-1.26%	-0.91%	-0.36%	-0.88%	-0.39%
$IR\text{-}only$	0.29%	1.84%	-1.56%	0.14%	0.15%
$No\text{-}risk$	-1.62%	-0.94%	-0.67%	-0.94%	-0.30%

Notes: GE: general equilibrium, PE: partial equilibrium, CO: crowding out; $AR\text{-}only$: economy with only aggregate risk, calibrated to match equity premium; $IR\text{-}only$: economy with only idiosyncratic risk; $No\text{-}risk$: deterministic economy. The total g_c is further decomposed into the mean effect, g_c^{mean}, and the distribution effect, g_c^{distr}, cf. Subsection 3.8 for formal definitions.

reduction of the total insurance gains. The reason is that aggregate wage volatility increases in this calibration which makes social security less attractive.

We also model an alternative distribution scheme, labelled BL-distr to L. Instead of distributing the contributions of workers to pensioners each period, we redistribute lump-sum to all workers. Such a scheme does not implement a life-time income risk smoothing like social security, but rather directly insures the idiosyncratic risk each period. We find that overall welfare gains are substantially smaller than in our baseline model. To understand this, observe that, on the one hand, such a redistributive scheme insures idiosyncratic risk in each period. As a consequence, the partial equilibrium gain from insuring idiosyncratic risk increases from 0.67% in the baseline model to 1.92%. On the other hand, this scheme does not provide any insurance against aggregate risk. Also, the partial equilibrium gain from insuring the CCV decreases from a pure insurance effect of 1.08% in the baseline model to 0.47%, because this redistribution scheme can only offer very limited insurance against the countercyclical variance: the volatility of idiosyncratic shocks is high in recessions, during which aggregate wages and the insurance payment to workers are low. This finding underscores once more the importance of risk interactions for the welfare benefits of social security and also gives support to the view that a lifetime redistributive scheme is desirable, see our concluding discussion.

Finally, instead of calibrating to the $\tau = 2\%$ economy we calibrate to the $\tau = 0\%$ economy, labelled scenario $BL_{\tau^{cal}=0\%}$. Targets remain the same as in our baseline calibration. We therefore have to give up some of the consistency of our baseline calibration strategy, because all our targets and empirical parameters are measured over a long period of time which always featured a social security system of moderate size. Since savings are higher in the $\tau = 0\%$ economy, the calibrated discount factor decreases to $\beta = 0.976$, which is about one percentage point lower than in our baseline calibration. As before, we target a risk-free return of 2.3%. The average stock return is at 3.0%. Both returns are therefore about one percentage point higher than in the $\tau = 0\%$ economy of our conservative baseline calibration. Results in Table 14 show that the introduction of social security still leads to welfare gains in general equilibrium, but gains now stand at about 0.5%, rather than the 2.5% we found for our baseline calibration. The share of welfare gains in partial equilibrium attributable to the risk interactions is at about 0.6, just as in our baseline calibration. One reason for the reduction in welfare gains relative to the baseline model is the higher asset returns in the calibrated $\tau = 0\%$ economy which makes social security a less attractive asset. This also causes small welfare losses of about -0.2% from the reduction of mean consumption. Another reason for the lower welfare gains is the lower value of the calibrated discount factor, which as we show in Corollary 1, is a crucial parameter for determining the welfare gains.

To illustrate this second mechanism, we also consider an experiment in which we hold constant the discount factor at the value of 0.987 resulting from the baseline calibration, cf. scenario $BL_{\tau^{cal}=0\% \,|\, \bar{\beta}}$. We accordingly give up targeting the capital output ratio, which now is at a still reasonable value of 3.3, as opposed to the baseline of 2.65 (cf. also scenario $BL\text{-}\frac{K}{Y} = 3$). Again, the risk-free return is calibrated to 2.3% and the risky stock return is about 3%. Relative to scenario $BL_{\tau^{cal}=0\%}$ welfare gains more than double (to 1.1%), while the share of interactions continues to stand at about 0.6. Welfare losses from the reduction of mean consumption remain at about -0.2%.

E.4 Life-Cycle Portfolios

Panel (a) of Figure 4 shows average shares invested in the risky asset over the life-cycle for model $EP_{IES=0.5}$, which comes closest to the calibration in standard life-cycle portfolio choice models like Cocco, Gomes, and Maenhout (2005). Recall from Section 3.2, in particular Footnote 15, that there is no borrowing constraint in our model. As a consequence, households are leveraged at the beginning of the life-cycle. From age 35 on, our simulated profiles are very similar to those presented in Cocco, Gomes, and Maenhout (2005).[60] Panel (b) shows corresponding life-cycle asset and stock holdings. Since there is no inter-generational transfer motive, asset holdings at young ages are very low and households fully decumulate assets towards the end of their life.

[60] Unlike Storesletten, Telmer, and Yaron (2007) we do not find in this model—and in any of our other scenarios—that the CCV mechanism leads to a hump-shaped portfolio share profile. We conjecture that this is a result of the perfect positive correlation between depreciation and technology shocks in their model, which maximizes the impact of CCV, so that young households hold less assets. We instead calibrate the correlation to the data, cf. Appendix D.3.

Table 14: Sensitivity Analysis: Other Scenarios

	g_c			$\frac{\Delta_{CCV}+\Delta_{CWG}}{g_c^{PE}}$	g_c^{distr}	g_c^{mean}
Scenario	GE	PE	CO		GE	
BL	+2.56%	+5.18%	-2.62%	0.60	+2.32%	+0.24%
BL-$\theta = 2$	+2.37%	+3.41%	-1.04%	0.53	+2.12%	+0.25%
BL-$\theta = 4$	+2.67%	+7.09%	-4.42%	0.65	+2.44%	+0.22%
BL-$\sigma_\varepsilon^2 = 0$	+2.44%	+4.30%	-1.86%	0.54	+2.22%	+0.22%
BL-$\bar{\kappa} = 0$	+2.08%	+4.46%	-2.38%	0.58	+1.82%	+0.27%
BL-$\frac{K}{Y} = 3$	+2.97%	+4.90%	-1.93%	0.60	+2.75%	+0.23%
BL-$\bar{\delta} = 0$	+1.19%	+2.62%	-1.42%	0.46	+0.78%	+0.42%
BL-distr to L	+1.61%	+1.92%	-0.31%	0.28	+1.57%	+0.03%
$BL_{\tau^{cal}=0\%}$	+0.47%	+3.39%	-2.92%	0.63	+0.67%	-0.19%
$BL_{\tau^{cal}=0\% \mid \bar{\beta}}$	+1.12%	+3.21%	-2.09%	0.61	+1.33%	-0.21%

Notes: GE: general equilibrium, PE: partial equilibrium, CO: crowding out; the total g_c is decomposed into the mean effect, g_c^{mean}, and the distribution effect, g_c^{distr}, cf. Subsection 3.8 for formal definitions.

Figure 4: Life-cycle Portfolios

(a) Risky Asset Share

(b) Assets and Stocks