# 7 Appendix

Appendix for

Capital Income Risk and the Dynamics of the Wealth Distribution

Hoang Khieu[(a)] and Klaus Wälde[(a,b)]
[(a)] Johannes Gutenberg University Mainz
[(b)] CESifo and IZA

September 26, 2018

# A The model

## A.1 The natural borrowing limit in (8)

When the individual is unemployed, they can borrow up to the present value of the flow of unemployment benefits minus minimum consumption. Using minimum consumption from (7), the present value of an unemployed worker that remains unemployed forever is given by $\int_t^\infty e^{-r[\tau-t]}\left(b(\tau)-c^{\min}\right)d\tau = \int_t^\infty e^{-r[\tau-t]}(1-\xi)\,b(\tau)\,d\tau$. Taking growth of the benefits from (3) into account, $b(\tau)=b(t)e^{g[\tau-t]}$, yields the smallest wealth level an individual can hold,

$$a^{\text{nat}}(t) = -(1-\xi)\,b(t)\int_t^\infty e^{-r[\tau-t]}e^{g[\tau-t]}d\tau = -\frac{(1-\xi)\,b(t)}{r-g}. \tag{A.1}$$

## A.2 Keynes-Ramsey rules in the model with trend

The four constraints in the main text (1), (2), (3) and (4) are reproduced here for convenience,

$$da(t) = \{r(t)\,a(t)+z(t)-c(t)\}\,dt, \tag{A.2}$$

$$dz(t) = gz(t)\,dt + [w(\Gamma(t))-z(t)]\,dq_\mu(t) + [b(\Gamma(t))-z(t)]\,dq_s(t), \tag{A.3}$$

$$\Gamma(t) = \Gamma_0 e^{gt}, \tag{A.4}$$

$$dr(t) = [r_{\text{high}}-r(t)]\,dq_{\text{low}}(t) + [r_{\text{low}}-r(t)]\,dq_{\text{high}}(t), \tag{A.5}$$

where $w(\Gamma(t)) = \hat{w}\Gamma(t)$ and $b(\Gamma(t)) = \hat{b}\Gamma(t)$. When we plug (A.4) into (A.3), we obtain a system of three state variables,

$$da(t) = \{r(t)\,a(t)+z(t)-c(t)\}\,dt, \tag{A.6}$$

$$dz(t) = gz(t)\,dt + \left[\hat{w}\Gamma_0 e^{gt}-z(t)\right]dq_\mu(t) + \left[\hat{b}\Gamma_0 e^{gt}-z(t)\right]dq_s(t), \tag{A.7}$$

$$dr(t) = [r_{\text{high}}-r(t)]\,dq_{\text{low}}(t) + [r_{\text{low}}-r(t)]\,dq_{\text{high}}(t). \tag{A.8}$$

These three variables, the wealth level $a(t)$, current income $z(t) \in \{w(\Gamma(t)), b(\Gamma(t))\}$ and the interest rate $r(t)$, describe the state $v(t) \equiv \{a(t), z(t), r(t)\}$ of an individual.

The individual maximizes their objective function by choosing a path $\{c(v(t))\}$ of consumption subject to the budget constraint (A.6) and the equation for their employment status (A.7). The changes in the interest rate in (A.8) are not anticipated as the individual is assumed to be myopic with respect to interest rate *changes* (for numerical reasons as discussed in footnote 23). Given the state $v(t)$, we define the value function as $V(v(t)) \equiv \max_{\{c(v(t))\}} U(t)$ subject to (A.6) and (A.7). The Bellman equation for this problem reads (see Sennewald, 2007, or Wälde, 2012, part IV)

$$\rho V(v(t)) = \max_{c(v(t))}\left\{u(c(v(t))) + \frac{1}{dt}E_t dV(v(t))\right\}. \tag{A.9}$$

Computing the differential $dV(v(t))$, taking the constraints (A.6) and (A.7) into account and forming expectations yields, suppressing time arguments for brevity,

$$\rho V(v) = \max_c \left\{ \begin{array}{c} u(c) + [ra + z - c] V_a(v) + gzV_z(v) \\ +s[V(a,b,r) - V(a,z,r)] + \mu [V(a,w,r) - V(a,z,r)] \end{array} \right\}, \tag{A.10}$$

where $V_x(v)$ stands for the partial derivative of $V(v)$ with respect to $x$, i.e. $V_x(v) \equiv \partial V(v)/\partial x$.[51] An optimal choice of consumption requires the first-order condition to equate marginal utility from consumption with the shadow price of wealth,

$$u'(c(v)) = V_a(v). \tag{A.11}$$

- Evolution of the shadow price

Using the budget constraint (A.2) and the evolution of labour income (A.3), the differential of the shadow price of wealth reads

$$\begin{aligned} dV_a(v) = {} & [V_{aa}(v)(ra + z - c) + gzV_{za}(v)] dt \\ & + [V_a(a,w,r) - V_a(a,z,r)] dq_\mu + [V_a(a,b,r) - V_a(a,z,r)] dq_s \end{aligned} \tag{A.12}$$

The maximized version of the Bellman equation (A.10) simply replaces the control variable $c$ by its optimal value $c(v)$,

$$\rho V(v) = \left\{ \begin{array}{c} u(c(v)) + [ra + z - c(v)] V_a(v) + gzV_z(v) \\ +s[V(a,b,r) - V(a,z,r)] + \mu [V(a,w,r) - V(a,z,r)] \end{array} \right\}. \tag{A.13}$$

Differentiating with respect to wealth yields, using the envelope theorem,

$$\rho V_a(v) = \left\{ \begin{array}{c} rV_a(v) + [ra + z - c(v)] V_{aa}(v) + gzV_{az}(v) \\ +s[V_a(a,b,r) - V_a(a,z,r)] + \mu [V_a(a,w,r) - V_a(a,z,r)] \end{array} \right\}. \tag{A.14}$$

After rearranging,

$$\begin{aligned} & (\rho - r) V_a(v) - s[V_a(a,b,r) - V_a(a,z,r)] - \mu [V_a(a,w,r) - V_a(a,z,r)] \\ & = [ra + z - c(v)] V_{aa}(v) + gzV_{az}(v). \end{aligned}$$

Inserting into (A.12), taking $V_{az}(v) = V_{za}(v)$ for granted,[52] gives

$$\begin{aligned} dV_a(v) = {} & \{(\rho - r) V_a(v) - s[V_a(a,b,r) - V_a(a,z,r)] - \mu [V_a(a,w,r) - V_a(a,z,r)]\} dt \\ & + [V_a(a,w,r) - V_a(a,z,r)] dq_\mu + [V_a(a,b,r) - V_a(a,z,r)] dq_s. \end{aligned} \tag{A.15}$$

- Inserting first-order condition

When we now replace the shadow price by marginal utility from the first-order condition (A.11), we get the Keynes-Ramsey rule for marginal utility,

$$\begin{aligned} du'(c(v)) = {} & \left\{ \begin{array}{c} (\rho - r) u'(c(v)) - s[u'(c(a,b,r)) - u'(c(a,z,r))] \\ -\mu [u'(c(a,w,r)) - u'(c(a,z,r))] \end{array} \right\} dt \\ & + [u'(c(a,w,r)) - u'(c(a,z,r))] dq_\mu + [u'(c(a,b,r)) - u'(c(a,z,r))] dq_s. \end{aligned} \tag{A.16}$$

[51] If individuals anticipated changes in interest rates from (A.8), two additional jump terms would appear in (A.10). This would lead to an additional reason for precautionary saving: When the interest rate is high, there is precautionary saving (as the interest rate could drop). When the interest rate is low, there is dissaving.

[52] We assume that the value function $V(\cdot)$ is continously increasing and concave and that its second partial derivatives are continuous. Thus, $V_{az}(v) = V_{za}(v)$ according to the Schwarz's theorem.

For an employed individual where, $z = w$, this reads

$$du'(c(a,w,r)) = \{(\rho - r)\, u'(c(a,z,r)) - s\,[u'(c(a,b,r)) - u'(c(a,w,r))]\}\, dt \\ + [u'(c(a,b,r)) - u'(c(a,w,r))]\, dq_s. \tag{A.17}$$

We now transform this optimality condition in marginal utilities into one in consumption levels. Let $f(.)$ be the inverse function for $u'$, i.e. $f(u') = c$ and apply the CVF to $f(u'(c(a,w,r)))$ This gives

$$df(u'(c(a,w,r))) = f'(u'(c(a,w,r)))\,\{(\rho - r)\, u'(c(a,w,r)) - s\,[u'(c(a,b,r)) - u'(c(a,w,r))]\}\, dt \\ + [f(u'(c(a,b,r))) - f(u'(c(a,w,r)))]\, dq_s.$$

As $f(u') = c$ and therefore $f'(u'(c(a,w,r))) = \frac{df(u'(c(a,w,r)))}{du'(c(a,w,r))} = \frac{dc(a,w,r)}{du'(c(a,w,r))} = \frac{1}{u''(c(a,w,r))}$, we get

$$dc(a,w,r) = \frac{1}{u''(c(a,w,r))}\{(\rho - r)\, u'(c(a,w,r)) - s\,[u'(c(a,b,r)) - u'(c(a,w,r))]\}\, dt \\ + [c(a,b,r) - c(a,w,r)]\, dq_s$$

which is equivalent to

$$\frac{u''(c(a,w,r))}{u'(c(a,w,r))} dc(a,w,r) = \left\{\rho - r - s\left[\frac{u'(c(a,b,r))}{u'(c(a,w,r))} - 1\right]\right\} dt \\ + \frac{u''(c(a,w,r))}{u'(c(a,w,r))}\,[c(a,b,r) - c(a,w,r)]\, dq_s. \tag{A.18}$$

Multiplying by "$-1$" and using the instantaneous CRRA utility function $u(c(t)) = \frac{c(t)^{1-\sigma}-1}{1-\sigma}$, we get $\frac{u''(c(a,w,r))}{u'(c(a,w,r))} = \frac{-\sigma}{c(a,w,r)}$ and after some rearrangements

$$dc(a,w,r) = \frac{c(a,w,r)}{\sigma}\left\{r - \rho + s\left[\left(\frac{c(a,w,r)}{c(a,b,r)}\right)^{\sigma} - 1\right]\right\} dt + [c(a,b,r) - c(a,w,r)]\, dq_s \tag{A.19}$$

which is (9a) in the main text.

The derivation of $dc(a,b,r)$ also starts from (A.16) and steps are in perfect analogy. Let us now use the notation $c_r(a,z)$ instead of $c(a,z,r)$ for convenience.

# B Detrending and equilibrium

## B.1 Evolution of detrended variables

Given the trend introduced in (3), detrended income is $\hat{z}(t) \equiv z(t)/\Gamma(t)$. Given the evolution of $z(t)$ from (4) and the fact that the trend $\Gamma(t)$ is deterministic, we obtain a SDE for $\hat{z}(t)$,

$$d\hat{z}(t) = d\frac{z(t)}{\Gamma(t)} = \left[\frac{1}{\Gamma(t)}(w(\Gamma(t)) - z(t))\right] dq_\mu(t) - \left[\frac{1}{\Gamma(t)}(b(\Gamma(t)) - z(t))\right] dq_s(t) \\ = [\hat{w} - \hat{z}]\, dq_\mu(t) + \left[\hat{b} - \hat{z}\right] dq_s(t). \tag{B.1}$$

- The budget constraint

We now compute the evolution of detrended wealth $\hat{a}(t) \equiv a(t)/\Gamma(t)$,

$$d\frac{a(t)}{\Gamma(t)} = \frac{\Gamma(t)\,da(t) - a(t)\,d\Gamma(t)}{\Gamma^2(t)} = \frac{\Gamma(t)\,[[ra(t)+z(t)-c(t)]\,dt] - a(t)\,g\Gamma(t)\,dt}{\Gamma^2(t)}$$
$$= \frac{1}{\Gamma(t)}\,[ra(t)+z(t)-c(t)-a(t)\,g]\,dt = \frac{1}{\Gamma(t)}\,[(r-g)\,a(t)+z(t)-c(t)]\,dt.$$

We can write this also as expression in detrended variables only, i.e.

$$d\hat{a}(t) = \{(r-g)\,\hat{a}(t)+\hat{z}(t)-\hat{c}(t)\}\,dt. \tag{B.2}$$

This describes the evolution of detrended wealth $\hat{a}(t)$ as a function of detrended income $\hat{z}(t)$ and detrended consumption $\hat{c}(t)$.

- Detrended consumption

The evolution of detrended consumption follows

$$d\hat{c}_r^{\hat{z}}(\hat{a}) \equiv d\frac{c_r^z(a(t))}{\Gamma(t)} = \frac{\Gamma(t)\,dc_r^z(a(t)) - c_r^z(a(t))d\Gamma(t)}{\Gamma^2(t)}.$$

For an employed worker, using eq. (A.19), this reads

$$\begin{aligned}
d\hat{c}_r^{\hat{w}}(\hat{a}(t)) &= d\frac{c_r(a(t),w(t))}{\Gamma(t)}\\
&= \frac{\Gamma(t)\left\{r-\rho+s\left[\left(\frac{c_r(a(t),w(t))}{c_r(a(t),b(t))}\right)^{\sigma}-1\right]\right\}\frac{c_r(a(t),w(t))}{\sigma} - c_r(a(t),w(t))\,g\Gamma(t)}{\Gamma^2(t)}dt\\
&= \frac{\left\{r-\rho+s\left[\left(\frac{c_r(a(t),w(t))}{c_r(a(t),b(t))}\right)^{\sigma}-1\right]\right\}\frac{c_r(a(t),w(t))}{\sigma} - c_r(a(t),w(t))\,g}{\Gamma(t)}dt\\
&= \frac{r-\rho-\sigma g+s\left[\left(\frac{c_r(a(t),w(t))}{c_r(a(t),b(t))}\right)^{\sigma}-1\right]}{\sigma}\frac{c_r(a(t),w(t))}{\Gamma(t)}dt\\
&= \left\{\frac{r-\rho}{\sigma}-g+\frac{s}{\sigma}\left[\left(\frac{c_r(a(t),w(t))\,\Gamma(t)^{-1}}{c_r(a(t),b(t))\,\Gamma(t)^{-1}}\right)^{\sigma}-1\right]\right\}\hat{c}_r^{\hat{w}}(\hat{a}(t))\,dt.
\end{aligned}$$

Employing the $\hat{c}_r^{\hat{z}}(\hat{a})$ notation, we obtain the detrended Keynes-Ramsey rule for optimal consumption of an employed worker,

$$d\hat{c}_r^{\hat{w}}(\hat{a}(t)) = \left\{\frac{r-\rho}{\sigma}-g+\frac{s}{\sigma}\left[\left(\frac{\hat{c}_r^{\hat{w}}(\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))}\right)^{\sigma}-1\right]\right\}\hat{c}_r^{\hat{w}}(\hat{a}(t))\,dt. \tag{B.3}$$

This result holds in a similar fashion for the case of an unemployed individual. We find

$$d\hat{c}_r^{\hat{b}}(\hat{a}(t)) = \left\{\frac{r-\rho}{\sigma}-g-\frac{\mu}{\sigma}\left[1-\left(\frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))}\right)^{\sigma}\right]\right\}\hat{c}_r^{\hat{b}}(\hat{a}(t))\,dt \tag{B.4}$$

## B.2 Consumption and wealth dynamics – low-interest-rate regime (15)

**Lemma 1** *In the low-interest-rate regime, i.e. when $r < \rho + \sigma g$ as in (15), the zero motion line for consumption when unemployed does not exist as $d\hat{c}_r^{\hat{b}}(\hat{a}(t))/dt < 0$ always holds.*

**Proof.** Consumption when unemployed falls iff

$$d\hat{c}_r^{\hat{b}}(\hat{a}(t))/dt < 0 \Leftrightarrow \frac{r-\rho}{\sigma} - g - \frac{\mu}{\sigma}\left[1 - \left(\frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))}\right)^{\sigma}\right] < 0$$

$$\Leftrightarrow r - \rho - \sigma g - \mu\left[1 - \left(\frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))}\right)^{\sigma}\right] < 0. \tag{B.5}$$

As $r < \rho + \sigma g$ in the low-interest-rate regime and $\hat{c}_r^{\hat{b}}(\hat{a}(t)) < \hat{c}_r^{\hat{w}}(\hat{a}(t))$, condition (B.5) always holds. ■

This lemma shows why condition (15) defines the low-interest-rate regime. Further properties of the low-interest-rate regime are as follows. The zero motion lines for wealth are given by

$$d\hat{a}_{\hat{w}}(t)/dt = 0 \Leftrightarrow \hat{c}_r^{\hat{w}}(t) = (r-g)\hat{a}(t) + \hat{w},$$
$$d\hat{a}_{\hat{b}}(t)/dt = 0 \Leftrightarrow \hat{c}_r^{\hat{b}}(t) = (r-g)\hat{a}(t) + \hat{b}.$$

Wealth when employed falls/rises over time iff $\hat{c}_r^{\hat{w}}(t) \gtrless (r-g)\hat{a}(t) + \hat{w}$. Similarly, wealth when unemployed falls/rises over time iff $\hat{c}_r^{\hat{b}}(t) \gtrless (r-g)\hat{a}(t) + \hat{b}$.

The zero motion line for consumption when employed, $\hat{c}_r^{\hat{w}}(\hat{a}(t))$, is given by

$$\frac{d\hat{c}_r^{\hat{w}}(\hat{a}(t))}{dt} = 0 \Leftrightarrow \frac{r-\rho}{\sigma} - g + \frac{s}{\sigma}\left[\left(\frac{\hat{c}_r^{\hat{w}}\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))}\right)^{\sigma} - 1\right] = 0$$

$$\Leftrightarrow \frac{\hat{c}_r^{\hat{w}}(\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))} = \left(1 - \frac{r-\rho-\sigma g}{s}\right)^{\frac{1}{\sigma}} \equiv \hat{\psi}^{-1}. \tag{B.6}$$

Consumption when employed falls iff

$$\frac{d\hat{c}_r^{\hat{w}}(\hat{a}(t))}{dt} < 0 \Leftrightarrow \frac{r-\rho}{\sigma} - g + \frac{s}{\sigma}\left[\left(\frac{\hat{c}_r^{\hat{w}}(\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))}\right)^{\sigma} - 1\right] < 0$$

$$\Leftrightarrow r - \rho - \sigma g + s\left[\left(\frac{\hat{c}_r^{\hat{w}}(\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))}\right)^{\sigma} - 1\right] < 0 \tag{B.7}$$

$$\Leftrightarrow \frac{\hat{c}_r^{\hat{w}}(\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))} < \left(1 - \frac{r-\rho-\sigma g}{s}\right)^{\frac{1}{\sigma}} = \hat{\psi}^{-1}. \tag{B.8}$$

When we express the change in detrended consumption in response to a change in detrended wealth using (B.2), (B.3), and (B.4), it yields our final two-dimensional ODE system for optimal consumption

$$\frac{d\hat{c}_r^{\hat{w}}(\hat{a})}{d\hat{a}} = \frac{\left\{\frac{r-\rho}{\sigma} - g + \frac{s}{\sigma}\left[\left(\frac{\hat{c}_r^{\hat{w}}(\hat{a})}{\hat{c}_r^{\hat{b}}(\hat{a})}\right)^{\sigma} - 1\right]\right\}\hat{c}_r^{\hat{w}}(\hat{a})}{(r-g)\hat{a} + \hat{w} - \hat{c}_r^{\hat{w}}(\hat{a})}, \tag{B.9a}$$

$$\frac{d\hat{c}_r^{\hat{b}}(\hat{a})}{d\hat{a}} = \frac{\left\{\frac{r-\rho}{\sigma} - g - \frac{\mu}{\sigma}\left[1 - \left(\frac{\hat{c}_r^{\hat{b}}(\hat{a})}{\hat{c}_r^{\hat{w}}(\hat{a})}\right)^{\sigma}\right]\right\}\hat{c}_r^{\hat{b}}(\hat{a})}{(r-g)\hat{a} + \hat{b} - \hat{c}_r^{\hat{b}}(\hat{a})}. \tag{B.9b}$$

For this ODE system to have a unique solution, we need two boundary conditions, one for $\hat{c}_r^{\hat{w}}(\hat{a})$ and the other for $\hat{c}_r^{\hat{b}}(\hat{a})$. (B.6) implies that there exists a $\hat{a}_{\hat{w}}^*$ that holds (B.6). When $\hat{a}_{\hat{w}}^*$ is unique, we can assume two boundary conditions to have a unique solution to the ODE system above. Thus, the zero motion line for consumption when employed is $\hat{a}(t) = \hat{a}_{\hat{w}}^*$, at which $\hat{c}_r^{\hat{b}}(\hat{a}_{\hat{w}}^*) = \hat{\psi}\hat{c}_r^{\hat{w}}(\hat{a}_{\hat{w}}^*)$. As we assume $d\hat{c}_r^{\hat{z}}(\hat{a})/d\hat{a} > 0$, consumption and wealth dynamics can be illustrated by fig. 1 in the main text.[53]

## B.3 Consumption and wealth dynamics – high-interest-rate regime (16)

**Lemma 2** *In the high-interest-rate regime, i.e when $r > \rho + \sigma g$ as in (16), the zero motion line for consumption when employed does not exist as $d\hat{c}_r^{\hat{w}}(\hat{a}(t))/dt > 0$ always holds.*

**Proof.** Consumption when employed falls iff

$$\begin{aligned} d\hat{c}_r^{\hat{w}}(\hat{a}(t))/dt > 0 &\Leftrightarrow \frac{r-\rho}{\sigma} - g + \frac{s}{\sigma}\left[\left(\frac{\hat{c}_r^{\hat{w}}(\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))}\right)^{\sigma} - 1\right] > 0 \\ &\Leftrightarrow r - \rho - \sigma g > -s\left[\left(\frac{\hat{c}_r^{\hat{w}}(\hat{a}(t))}{\hat{c}_r^{\hat{b}}(\hat{a}(t))}\right)^{\sigma} - 1\right]. \end{aligned} \tag{B.10}$$

As $r > \rho + \sigma g$ in the high-interest-rate regime and $\hat{c}_r^{\hat{b}}(\hat{a}(t)) < \hat{c}_r^{\hat{w}}(\hat{a}(t))$, condition (B.10) always holds. ■

This is why (16) defines the high-interest-rate regime. Further properties of the high-interest-rate regime are as follows. The zero motion lines for wealth and wealth dynamics are of perfect analogy as in the low-interest-rate regime.

The zero motion line for consumption when unemployed, $\hat{c}_r^{\hat{b}}(\hat{a}(t))$, is given by

$$\begin{aligned} \frac{d\hat{c}_r^{\hat{b}}(\hat{a}(t))}{dt} = 0 &\Leftrightarrow \frac{r-\rho}{\sigma} - g - \frac{\mu}{\sigma}\left[\left(1 - \frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))}\right)^{\sigma}\right] = 0 \\ &\Leftrightarrow \frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))} = \left(1 - \frac{r-\rho-\sigma g}{\mu}\right)^{\frac{1}{\sigma}} \equiv \hat{\psi}_{\hat{b}}. \end{aligned} \tag{B.11}$$

Consumption when unemployed rises iff

$$\frac{d\hat{c}_r^{\hat{b}}(\hat{a}(t))}{dt} = 0 > 0 \Leftrightarrow \frac{r-\rho}{\sigma} - g - \frac{\mu}{\sigma}\left[\left(1 - \frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))}\right)^{\sigma}\right] > 0$$

$$\Leftrightarrow r - \rho - \sigma g - \mu\left[\left(1 - \frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))}\right)^{\sigma}\right] > 0 \tag{B.12}$$

$$\Leftrightarrow \frac{\hat{c}_r^{\hat{b}}(\hat{a}(t))}{\hat{c}_r^{\hat{w}}(\hat{a}(t))} > \left(1 - \frac{r-\rho-\sigma g}{\mu}\right)^{\frac{1}{\sigma}}. \tag{B.13}$$

Using the same logic as in the low-interest-rate regime, we argue that there exists a $\hat{a}_{\hat{b}}^*$ such that (B.11) holds. Thus, the zero motion line for consumption when unemployed is $\hat{a}(t) = \hat{a}_{\hat{b}}^*$, at which $\hat{c}_r^{\hat{b}}\left(\hat{a}_{\hat{b}}^*\right) = \hat{\psi}_{\hat{b}}\hat{c}_r^{\hat{w}}\left(\hat{a}_{\hat{b}}^*\right)$. As we assume $d\hat{c}_r^{\hat{z}}(\hat{a})/d\hat{a} > 0$, consumption and wealth dynamics can be illustrated as in fig. 2 in the main text.

[53] Proofs for these intuitive properties can be found in Bayer and Wälde (2010, 2015).

# C Distributional dynamics - deriving Fokker-Planck equations

We derive equations (23a) and (23b) from the main text following five steps below. We use the approach described in Bayer and Wälde (2010a, sect. 5) to derive the Fokker-Planck equations for our case.

- The expected change of some function $f$

Assume there is some arbitrary function $f$ having the state variables $\hat{a}$ and $\hat{z}$ as arguments. The stochastic processes $\hat{a}$ and $\hat{z}$ are respectively described as follows

$$d\hat{a}(\tau) = \{(r-g)\hat{a}(\tau) + \hat{z}(\tau) - \hat{c}(\tau)\} d\tau, \tag{C.1}$$

$$d\hat{z}(\tau) = [\hat{w} - \hat{z}] dq_\mu(\tau) + \left[\hat{b} - \hat{z}\right] dq_s(\tau), \tag{C.2}$$

where we use the notation $\hat{c}(\tau)$ instead of $\hat{c}^{\hat{z}}_\tau(\tau)$ for convenience. Function $f$ has a bounded support $S = (\hat{a}_{\min}, \hat{a}_{\max}) \times \{\hat{b}, \hat{w}\}$, i.e. $f(\hat{a}, \hat{z}) = 0$ for $(\hat{a}, \hat{z}) \notin S$. The differential of this function reads (see (Wälde 2012), ch. 10)

$$\begin{aligned} df(\hat{a}(\tau), \hat{z}(\tau)) = & \frac{\partial}{\partial \hat{a}} f(\hat{a}(\tau), \hat{z}(\tau)) \left[(r-g)\hat{a}(\tau) + \hat{z}(\tau) - \hat{c}(\tau)\right] d\tau \\ & + \left[f(\hat{a}(\tau), \hat{w}) - f(\hat{a}(\tau), \hat{z}(\tau))\right] dq_\mu(\tau) \\ & + \left[f\left(\hat{a}(\tau), \hat{b}\right) - f(\hat{a}(\tau), \hat{z}(\tau))\right] dq_s(\tau). \end{aligned}$$

Applying the conditional expectations operator $\mathbb{E}_t$ and dividing by $dt$ yields the heuristic equation

$$\begin{aligned} \frac{\mathbb{E}_\tau df(\hat{a}(\tau), \hat{z}(\tau))}{d\tau} = & \frac{\partial}{\partial \hat{a}} f(\hat{a}(\tau), \hat{z}(\tau)) \left[(r-g)\hat{a}(\tau) + \hat{z}(\tau) - \hat{c}(\tau)\right] \\ & + \mu \left[f(\hat{a}(\tau), \hat{w}) - f(\hat{a}(\tau), \hat{z}(\tau))\right] \\ & + s\left[f\left(\hat{a}(\tau), \hat{b}\right) - f(\hat{a}(\tau), \hat{z}(\tau))\right]. \end{aligned}$$

In what follows, we denote this expression by

$$\mathcal{A} f(\hat{a}(\tau), \hat{z}(\tau)) \equiv \frac{\mathbb{E}_\tau df(\hat{a}(\tau), \hat{z}(\tau))}{d\tau}. \tag{C.3}$$

- Dynkin's formula and its manipulation

To abbreviate notation, we now define $x(\tau) := (\hat{a}(\tau), \hat{z}(\tau))$. The expected value of our function $f(x(\tau))$ is, by Dynkin's formula (see e.g. Yuan and Mao, 2003), given by

$$\mathbb{E} f(x(\tau)) = \mathbb{E} f(x(t)) + \int_t^\tau \mathbb{E}\left(\mathcal{A} f(x(s))\right) ds. \tag{C.4}$$

To understand this equation, use the definition in (C.3) and formally write it as

$$\mathbb{E} f(x(\tau)) = \mathbb{E} f(x(t)) + \int_t^\tau \frac{\mathbb{E} df(x(s))}{ds} ds = \mathbb{E} f(x(t)) + \int_t^\tau \mathbb{E} df(x(s)).$$

This equation implies that the (unconditional) expected value of $f(x)$ at any point in time $\tau$ is equal to the expectation for the current value, $\mathbb{E} f(x(t))$ (given that x(t) is random), plus the "sum" of the expected future changes from time $t$ to $\tau$, $\int_t^\tau \mathbb{E} df(x(s)) ds$.

Let us now differentiate (C.4) with respect to $\tau$, i.e. we now ask how the expectations about $f(x(\tau))$ change when $\tau$ moves further into the future. We find that

$$\frac{\partial}{\partial \tau}\mathbb{E}f(x(\tau)) = \frac{\partial}{\partial \tau}\int_t^{\tau} \mathbb{E}\left(\mathcal{A}f(x(s))\right) ds = \mathbb{E}\left(\mathcal{A}f(x(\tau))\right), \tag{C.5}$$

where the first equality used that $\mathbb{E}f(x(t))$ is a constant, and the second equality used the Leibniz rule. Equation (C.5) says that the change in the expectations about $f(x(\tau))$ is equal to the expected change of $f(x(\tau))$, where the change is $\mathcal{A}f(x(\tau))$. This is equation (14.11) in Davis (1993), the "abstract version of the <...> Kolmogorov backward equation" (p. 30).

We now introduce the density $p(\hat{a},\hat{z},\tau)$. The expectation operator $\mathbb{E}$ in (C.5) integrates over all possible states of $x(\tau)$. At each point in time, $\hat{a}$ is continuous while $\hat{z}$ is discrete and $\hat{z} \in \{\hat{b},\hat{w}\}$. Thus, we can express this joint density as

$$\begin{aligned} p(\hat{a},\hat{z},\tau) &= p(\hat{a},\tau|\hat{z})p_{\hat{z}}(\tau) = p(\hat{a},\tau|\hat{w})p_{\hat{w}}(\tau) + p(\hat{a},\tau|\hat{b})p_{\hat{b}}(\tau) \\ &= p(\hat{a},\hat{w},\tau) + p(\hat{a},\hat{b},\tau) = p^{\hat{w}}(\hat{a},\tau) + p^{\hat{b}}(\hat{a},\tau) \end{aligned}$$

where $p_{\hat{z}}(\tau)$ denotes the probability of being in employment–state $\hat{z}$ at time $\tau$. As $p^{\hat{w}}(\hat{a},\tau)$ and $p^{\hat{b}}(\hat{a},\tau)$ are densities of (detrended) wealth and time, we then can express

$$p(\hat{a},\tau) = p^{\hat{w}}(\hat{a},\tau) + p^{\hat{b}}(\hat{a},\tau),$$

i.e. the joint density of (detrended) wealth and time is equal to a "sub-density" of (detrended) wealth and time when $\hat{z}=\hat{w}$ plus a "sub-density" of (detrended) wealth and time when $\hat{z}=\hat{b}$. We now can write (C.5) as

$$\begin{aligned} \frac{\partial}{\partial \tau}\mathbb{E}f(x(\tau)) &= \mathbb{E}\left(\mathcal{A}f(x(s))\right) \\ &= p_{\hat{w}}(\tau)\int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mathcal{A}f(\hat{a},\hat{w})p(\hat{a},\tau|\hat{w})d\hat{a} + p_{\hat{b}}(\tau)\int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mathcal{A}f(\hat{a},\hat{b})p(\hat{a},\tau|\hat{b})d\hat{a} \\ &= \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mathcal{A}f(\hat{a},\hat{w})p(\hat{a},\hat{w},\tau)d\hat{a} + \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mathcal{A}f(\hat{a},\hat{b})p(\hat{a},\hat{b},\tau)d\hat{a} \\ &= \Phi_{\hat{w}} + \Phi_{\hat{b}} \end{aligned} \tag{C.6}$$

where $\Phi_{\hat{w}} := \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mathcal{A}f(\hat{a},\hat{w})p(\hat{a},\hat{w},\tau)d\hat{a}$ and $\Phi_{\hat{b}} := \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mathcal{A}f(\hat{a},\hat{b})p(\hat{a},\hat{b},\tau)d\hat{a}$.

- The adjoint operator and integration by parts

This is now the crucial step in obtaining a differential equation for the density. It consists of applying an integration by parts formula which allows to move the derivatives in $\mathcal{A}f(x(\tau))$ into the density $p(x,\tau)$. Let us briefly review this method, without getting into technical details. Given two functions $f,g:\mathbb{R}\to\mathbb{R}$ and two fixed real numbers $c<d$, the product rule of differentiation

$$d(f(x)g(x)) = df(x)g(x) + f(x)dg(x) \tag{C.7}$$

implies that $f(d)g(d) - f(c)g(c) = \int_c^d f'(x)g(x)dx + \int_c^d f(x)g'(x)dx$, a formula referred to as partial integration rule. In particular, it also holds for $c=-\infty$ and $d=+\infty$, if the function evaluations are understood as limits for $c\to-\infty$ and $d\to+\infty$, respectively. If $f$ has bounded support, i.e. is equal to zero outside a fixed bounded set, then the function evaluations at $\pm\infty$ vanish and we get

$$\int_{-\infty}^{+\infty} f'(x)g(x)dx = -\int_{-\infty}^{+\infty} f(x)g'(x)dx \tag{C.8}$$

We now apply (C.8) to (C.6). But we first compute

$$\begin{aligned}
\mathcal{A}f(\hat{a},\hat{w}) =& \frac{\partial}{\partial \hat{a}} f(\hat{a},\hat{w}) \left((r-g)\hat{a} + \hat{w} - \hat{c}_r^{\hat{w}}(\hat{a})\right) + s\left(f\left(\hat{a},\hat{b}\right) - f(\hat{a},\hat{w})\right) \\
\mathcal{A}f(\hat{a},\hat{b}) =& \frac{\partial}{\partial \hat{a}} f(\hat{a},\hat{b}) \left((r-g)\hat{a} + \hat{b} - \hat{c}_r^{\hat{b}}(\hat{a})\right) + \mu\left(f\left(\hat{a},\hat{w}\right) - f(\hat{a},\hat{b})\right),
\end{aligned}$$

Now applying integration by parts yields

$$\begin{aligned}
&\Phi_{\hat{w}} = \\
&- \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{w}) \left\{ \left[ r - g - \frac{d\hat{c}_r^{\hat{w}}(\hat{a})}{d\hat{a}} \right] p(\hat{a},\hat{w},\tau) + \left[(r-g)\hat{a} + \hat{w} - \hat{c}_r^{\hat{w}}(\hat{a})\right] \frac{\partial}{\partial \hat{a}} p(\hat{a},\hat{w},\tau) \right\} d\hat{a} \\
&+ \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} s\left(f\left(\hat{a},\hat{b}\right) - f(\hat{a},\hat{w})\right) p(\hat{a},\hat{w},\tau) d\hat{a}
\end{aligned} \tag{C.9}$$

$$\begin{aligned}
&\Phi_{\hat{b}} = \\
&- \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{b}) \left\{ \left[ r - g - \frac{d\hat{c}_r^{\hat{b}}(\hat{a})}{d\hat{a}} \right] p(\hat{a},\hat{b},\tau) + \left[(r-g)\hat{a} + \hat{b} - \hat{c}_r^{\hat{b}}(\hat{a})\right] \frac{\partial}{\partial \hat{a}} p(\hat{a},\hat{b},\tau) \right\} d\hat{a} \\
&+ \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mu\left(f\left(\hat{a},\hat{w}\right) - f(\hat{a},\hat{b})\right) p(\hat{a},\hat{b},\tau) d\hat{a}
\end{aligned} \tag{C.10}$$

Thus, we find

$$\begin{aligned}
&\frac{\partial}{\partial \tau} \mathbb{E} f(x(\tau)) = \Phi_{\hat{w}} + \Phi_{\hat{b}} = \\
&- \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{w}) \left\{ \left[ r - g - \frac{d\hat{c}_r^{\hat{w}}(\hat{a})}{d\hat{a}} \right] p(\hat{a},\hat{w},\tau) + \left[(r-g)\hat{a} + \hat{w} - \hat{c}_r^{\hat{w}}(\hat{a})\right] \frac{\partial}{\partial \hat{a}} p(\hat{a},\hat{w},\tau) \right\} d\hat{a} \\
&+ \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} s\left(f\left(\hat{a},\hat{b}\right) - f(\hat{a},\hat{w})\right) p(\hat{a},\hat{w},\tau) d\hat{a} \\
&- \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{b}) \left\{ \left[ r - g - \frac{d\hat{c}_r^{\hat{b}}(\hat{a})}{d\hat{a}} \right] p(\hat{a},\hat{b},\tau) + \left[(r-g)\hat{a} + \hat{b} - \hat{c}_r^{\hat{b}}(\hat{a})\right] \frac{\partial}{\partial \hat{a}} p(\hat{a},\hat{b},\tau) \right\} d\hat{a} \\
&+ \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} \mu\left(f\left(\hat{a},\hat{w}\right) - f(\hat{a},\hat{b})\right) p(\hat{a},\hat{b},\tau) d\hat{a}
\end{aligned} \tag{C.11}$$

- The expected value-an alternative to the Dynkin's formula

Let us now derive the second expression for the change in the expected value. By definition, and as an alternative to the Dynkin formula (C.4), we have

$$\mathbb{E} f(x(\tau)) = \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{w}) p(\hat{a},\hat{w},\tau) d\hat{a} + \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{b}) p(\hat{a},\hat{b},\tau) d\hat{a}$$

Differentiating this expression with respect to $\tau$ gives

$$\frac{\partial}{\partial \tau} \mathbb{E} f(x(\tau)) = \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{w}) \frac{\partial}{\partial \tau} p(\hat{a},\hat{w},\tau) d\hat{a} + \int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{b}) \frac{\partial}{\partial \tau} p(\hat{a},\hat{b},\tau) d\hat{a} \tag{C.12}$$

- Equating the two expressions

We now equate (C.11) and (C.12). Collecting terms belonging to $f(\hat{a},\hat{w})$ and $f(\hat{a},\hat{b})$ gives

$$\int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{w})\varphi_{\hat{w}}d\hat{a}+\int_{\hat{a}_{\min}}^{\hat{a}_{\max}} f(\hat{a},\hat{b})\varphi_{\hat{b}}d\hat{a}=0 \tag{C.13}$$

where

$$\begin{aligned}\varphi_{\hat{w}}=&-\frac{\partial}{\partial\tau}p(\hat{a},\hat{w},\tau)-\left[r-g-\frac{d\hat{c}_r^{\hat{w}}(\hat{a})}{d\hat{a}}+s\right]p(\hat{a},\hat{w},\tau)+\mu p(\hat{a},\hat{b},\tau)\\&-\left[(r-g)\hat{a}+\hat{w}-\hat{c}_r^{\hat{w}}(\hat{a})\right]\frac{\partial}{\partial\hat{a}}p(\hat{a},\hat{w},\tau)\end{aligned} \tag{C.14}$$

$$\begin{aligned}\varphi_{\hat{b}}=&-\frac{\partial}{\partial\tau}p(\hat{a},\hat{b},\tau)-\left[r-g-\frac{d\hat{c}_r^{\hat{b}}(\hat{a})}{d\hat{a}}+\mu\right]p(\hat{a},\hat{b},\tau)+sp(\hat{a},\hat{w},\tau)\\&-\left[(r-g)\hat{a}+\hat{b}-\hat{c}_r^{\hat{b}}(\hat{a})\right]\frac{\partial}{\partial\hat{a}}p(\hat{a},\hat{b},\tau)\end{aligned} \tag{C.15}$$

As (C.13) holds for any arbitrary function $f$, it requires that

$$\varphi_{\hat{w}}=0 \tag{C.16}$$
$$\varphi_{\hat{b}}=0 \tag{C.17}$$

When we rearrange and evaluate (C.16) and (C.17) at $\tau=t$ we obtain

$$\frac{\partial}{\partial t}p^{\hat{w}}(\hat{a},t)+\left[(r-g)\hat{a}+\hat{w}-\hat{c}_r^{\hat{w}}(\hat{a})\right]\frac{\partial}{\partial\hat{a}}p^{\hat{w}}(\hat{a},t)=\left[\frac{dc_r^{\hat{w}}(\hat{a})}{d\hat{a}}-(r-g)-s\right]p^{\hat{w}}(\hat{a},t)+\mu p^{\hat{b}}(\hat{a},t), \tag{C.18}$$

$$\frac{\partial}{\partial t}p^{\hat{b}}(\hat{a},t)+\left[(r-g)\hat{a}+\hat{b}-\hat{c}_r^{\hat{b}}(\hat{a})\right]\frac{\partial}{\partial\hat{a}}p^{\hat{b}}(\hat{a},t)=sp^{\hat{w}}(\hat{a},t)+\left[\frac{dc_r^{\hat{b}}(\hat{a})}{d\hat{a}}-(r-g)-\mu\right]p^{\hat{b}}(\hat{a},t), \tag{C.19}$$

where $p^{\hat{w}}(\hat{a},t)\equiv p(\hat{a},\hat{w},t)$ and $p^{\hat{b}}(\hat{a},t)\equiv p(\hat{a},\hat{b},t)$. These are two Fokker–Planck equations.

# D The empirical fit

## D.1 Data and quantitative phase diagram

### D.1.1 Computing continuous time variables

- The interest rate

Imagine we observe a yearly interest rate of 4%. The relationship between a monthly interest rate $r_m$ and a yearly rental rate $r_y$ is given by

$$a_0\left[1+r_y\right]=a_0\left[1+r_m\right]^{12}\Leftrightarrow r_m=\sqrt[12]{1+r_y}-1. \tag{D.1}$$

We can therefore easily compute the monthly interest rate. If we want the continuous time analogue to the monthly interest rate, we start from the equation that tells us how a certain capital stock evolves over time with continuous interest payments. This equation reads $\dot{a}=ra$

whose solution is of course $a(t) = a_0 e^{rt}$. We then obtain the continuous time interest rate by using

$$a_0 e^r = a_0 [1 + r_m]^n \tag{D.2}$$

where the unit of time is implicitly defined by this equation to mean $n$ months. The continuous time interest rate is then given by

$$r = n \ln [1 + r_m]. \tag{D.3}$$

As the unit of time in our model is 1 year, $n = 12$ here and in what follows.

- The time preference rate

We can undertake the same steps and go from a continuous time preference rate to a corresponding discount factor. When we replace $r$ by $\rho$ in (D.2) (where we should use 'utility' for interpretation purposes rather than wealth) and solve for the monthly time preference rate, we get $a_0 e^\rho = a_0 [1 + \rho_m]^n \Leftrightarrow e^{\rho/n} - 1 = \rho_m$. The annual time preference rate is then from (D.1)

$$\rho_y = [1 + \rho_m]^{12} - 1 = e^{\rho 12/n} - 1. \tag{D.4}$$

- The wage rate and its unit

Now consider a, say, monthly paid wage. In order to compute the corresponding wage *rate* in a continuous time model, we start from an equation similar to the starting point of (D.2) or similar to a budget constraint, $\dot{a} = ra + w$. The solution to this equation reads $a(t) = a_0 e^{rt} + \frac{w}{r}[e^{rt} - 1]$. We now use an equation in spirit similar to (D.2),

$$a_0 e^r + \frac{w}{r}[e^r - 1] = a_0 [1 + r_m]^n + \Sigma_{i=1}^n \frac{w_m}{\Phi} [1 + r_m]^{n-i}. \tag{D.5}$$

The left hand side is the wealth level at $t = 1$ with a continuous interest rate of $r$ and a wage rate of $w$. The right hand side is the wealth level at a monthly interest rate and a monthly wage of $w_m$, where the latter is scaled by a constant $\Phi$. Given that (D.2) holds, this equation simplifies to

$$\frac{w}{r}[e^r - 1] = \Sigma_{i=1}^n \frac{w_m}{\Phi} [1 + r_m]^{n-i} \Leftrightarrow w = \frac{r}{e^r - 1} \Sigma_{i=1}^n \frac{w_m}{\Phi} [1 + r_m]^{n-i}. \tag{D.6}$$

Concerning units of measurement, the monthly wage $w_m$ is measured in US\$. Then the unit of measurement in the model is US\$ as well. If we scale the monthly wage by $\Phi = 1000$ or by $\Phi = 345.07$, the wage rate is measured in 1000 US\$ or US\$. In fact, we set $\Phi = 1000$ for our numerical solution. The same logic applies to other payments like unemployment benefits or pension payments or other.

- Wage growth

Let the annual growth of the real wage be $g^w$ (which is 6.5% in our sample). Then the continuous time counterpart follows from

$$\hat{w}[1 + g^w] = \hat{w} e^{g12/n}.$$

The left hand side gives the wage after on year with an annual growth rate of $g^w$. When one unit of time in the continuous time model is $n$ months and a year has 12 months, the right hand side yields the continuous time growth rate $g$ as

$$\frac{n}{12} \ln [1 + g^w] = g. \tag{D.7}$$

If we go the other way round, we get

$$g^w = e^{g12/n} - 1.$$

- Arrival rates

Imagine we observe average durations in unemployment and in employment and we want to model durations by Poisson processes. Then the expected number of jumps over an interval of length $n$ (say months) is given by $\lambda n$ where $\lambda$ is the arrival rate of the Poisson process. If we fix one unit of time to $n$, the expected number of jumps per time interval is $\lambda$. The expected duration in a state is then given by $\lambda^{-1}$ units of time. If we observe an average duration of $d_m$ months in the data, we can use

$$d_m = \frac{n}{\lambda}. \tag{D.8}$$

### D.1.2 Parameter values

We describe here how parameters are chosen or estimated.

- The interest rate

We fix the interest rate $r$ such that it corresponds to an annual interest rate of 4%. We translate this into a continuous time interest rate $r$ by $e^{r12/n} = 1.04$, where $n$ is the number of months one unit of time in continuous time stands for. This yields $r = \frac{n}{12} \ln 1.04 = 0.0033n$. With a lower interest rate of 3%, we would end up at $r = 0.0025n$ and would obtain an infinite natural borrowing constraint given wage growth (see below) of $g = 0.0027n$.

- Income $w$ and $b$

The survey includes income from salary and wages in the previous year. We use this variable and adjust for the number of weeks worked in that specific year to compute annual income. As one unit of time in the model equals one year, we set $w$ equal to the annual income.

The value of $b$ cannot be determined from the NLSY because we lack information for later years. We therefore employ a replacement rate as discussed in the main text (see footnote 35).

- The wage growth rate $g$

The wage in the model grows as described before (3), i.e. $w(\Gamma(t)) = \hat{w}\Gamma_0 e^{gt}$. We equate $\hat{w}\Gamma_0$ to the mean wage of 1986, $w_{86}$ (expressed in 2008 $US\$$). The mean wage of 2008, $w_{08}$, then allows us to determine $g$ by solving $w\left(\Gamma\left(\frac{264}{n}\right)\right) = w_{86}e^{g\frac{264}{n}} = w_{08}$. (Our observation period are 264 months, i.e. 22 years). This yields $g = \ln\left(\frac{w_{08}}{w_{86}}\right)\frac{n}{264}$ and, given the data, we obtain $g = 0.0027n$.

- Job-arrival and separation rates, $\mu$ and $s$

The arrival rates are linked to the duration in each state via (D.8), i.e. $d_e = \frac{n}{s}$ and $d_u = \frac{n}{\mu}$. To determine an approximate value for the mean duration in employment and unemployment, we use the first cross-section and look at the weekly employment history of those individuals. We discard unemployment spells of less than 2 weeks as this is likely to be connected with job-to-job-transitions. With $n = 12$ months, we obtain the durations in table 1.

- The natural borrowing limit

Data provides us with some minimum wealth level $a^{\min}_{\text{data}}$. By equating $a^{\min}_{\text{data}}$ with the natural borrowing limit (A.1), we obtain

$$a^{\min}_{\text{data}} = -\frac{(1-\xi)\, b(t)}{r-g} \Leftrightarrow 1-\xi = -(r-g)\frac{a^{\min}_{\text{data}}}{b(t)}.$$

This provides us with an estimate of how much consumption is needed for survival. Plausibility requires $0 < \xi < 1$. Practically speaking, we choose a value $\hat{a}^{\text{nat}}(t)$ for the detrended model, i.e. for 1986 from which we start, and solve for optimal consumption and the wealth distribution for the range provided in (21).

## D.2 Targeting wealth distributions and measuring the fit

### D.2.1 The minimization problem for hitting wealth distributions

- One target year

Imagine we target 2008 and we need weights for the different interest rate paths. We obtain them by numerically solving the minimization problem

$$\min_{\{p_j\}} [1 - F(t)] \tag{D.9}$$
$$\text{subject to } p_j \geq 0 \forall j \text{ and } \Sigma_{j=1}^{n} p_j = 1.$$

The optimal number $n$ of financial types is obtained by computing the fit for $n \in \{2, ..., 130\}$ and by picking the maximum.

- Many target years

Now imagine we want to maximize the fit over years 1 to $T$. Then the minimization problem reads $\min_{\{p_j\}} \Sigma_{t=1}^{T} [1 - F(t)]$ subject to the same constraints as in (D.9). The optimal number of financial types is chosen as before.

### D.2.2 The density for wealth $a$

Following e.g. Wackerly (2008, ch. 6.4) or Wälde (2012, theorem 7.3.2), we have a random variable $x(t) = \hat{a}(t)$ with a density $f(x(t)) = p(\hat{a}, t)$ on a range $[\hat{a}_1(t), \hat{a}_2(t)]$ and a transformation of it which is $y(t) = y(x(t))$ where $y(t) = a(t)$ and $y(x(t)) = \hat{a}(t)\Gamma(t)$. The density $g(y) = g(a(t))$ is then given by $f(x(y)) \frac{dx}{dy}$ on the range $[y(\hat{a}_1(t)), y(\hat{a}_2(t))]$ function.

As $x(y) = a(t)/\Gamma(t)$, we have $dx/dy = d\hat{a}(t)/da(t) = 1/\Gamma(t)$. This means

$$g(a(t)) = f(x(y(t))) \frac{dx(t)}{dy(t)} = p\left(\frac{a(t)}{\Gamma(t)}, t\right) \frac{1}{\Gamma(t)}.$$

The support of $a(t)$ is $[\hat{a}_1(t)\Gamma(t), \hat{a}_2(t)\Gamma(t)]$.

### D.2.3 The quantitative support of wealth

The support of wealth for computing the evolution of distributions over time is stationary. Once we add the trend, the support is as illustrated in fig. 10.

We take the data and numerically construct a density from the empirical histogram. As the density has a much larger support than what is actually in the data, we cut 0.01% below (no adjustment is needed above). The cyan line in this figure shows the minimum and maximum of the support of the constructed density. After truncating below, we obtain the blue line at the lower end. (As we do not cut above, the blue and cyan lines coincide.) The red curve shows the minimum and maximum of wealth levels in the original survey. Finally, the black curve shows the model support of wealth when targeting 2008 density. Note that while the cyan, blue, and red curves do not change, the black curve does change according to different targeted years.

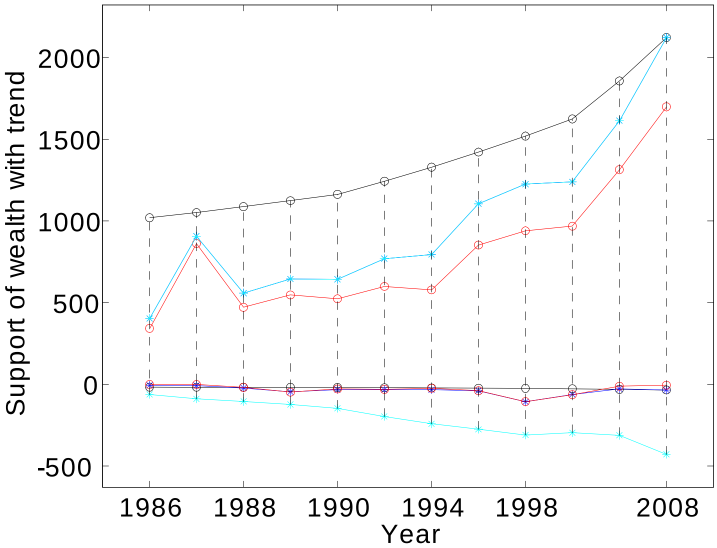


**Figure 10** *Supports of wealth when 2008 density is targeted*

| | 1986 | 1987 | 1988 | 1989 | 1990 | 1992 | 1994 | 1996 | 1998 | 2000 | 2004 | 2008 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| min | -16.3 | -16.9 | -17.4 | -18.0 | -18.6 | -19.9 | -21.3 | -22.8 | -24.3 | -26.0 | -29.7 | -34.0 |
| max | 1,020 | 1,052 | 1,088 | 1,125 | 1,163 | 1,244 | 1,330 | 1,421 | 1,520 | 1,625 | 1,857 | 2,123 |

**Table 5** *Minimum and maximum wealth levels (in 1000 US\$) in the data*

### D.2.4 The fit over 12 waves for different targets

We show densities for all 12 waves with wealth information for the case where we target one density (2008) and where we target the average fit over all years.

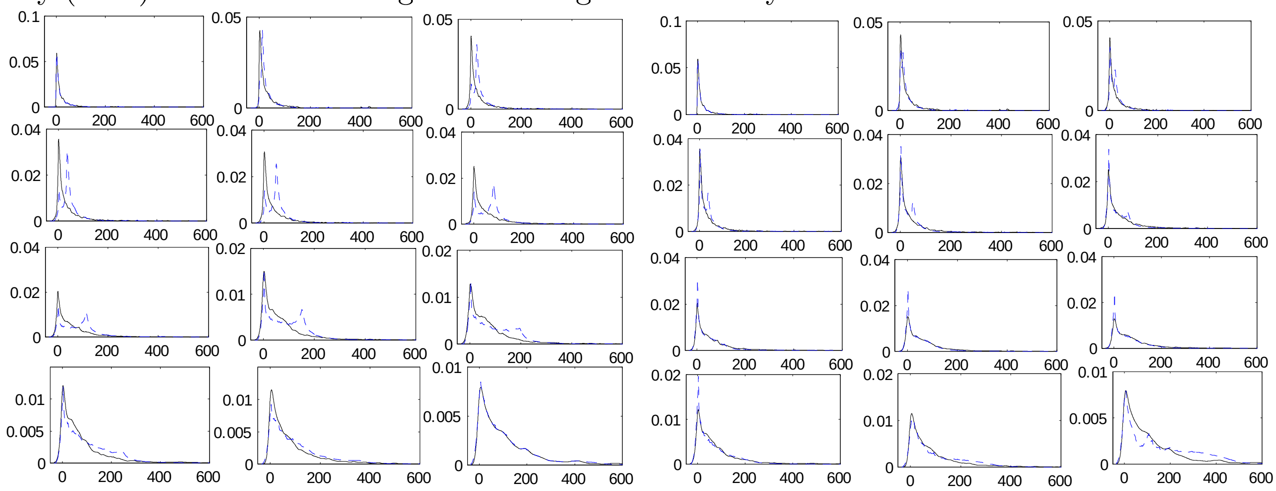

**Figure 11** *Fit for targeting 2008 (left part) and for targeting the average (right part)*

The left figure shows the densities for all years when 2008 is targeted. The right figure shows the same for targeting the average fit over all years. The empirical densities are in black while the model densities are in blue.

## D.3 Robustness checks

### D.3.1 Pure capital income risk with a two-point interest rate distribution

The analysis of pure capital income risk is simpler as individuals that are myopic with respect to interest rate changes and that do not face labour income uncertainty behave as if they lived in a deterministic world. This section therefore first describes the analytical approach and then goes to results.

- The analytical approach

As always, we employ $t$ for the current or initial point in time and $\tau$ for a future point in time. We solve an optimal consumption problem with an intertemporal objective function as in (5) in the main text. For analytical convenience, the instantaneous utility function (6) is replaced by a Stone-Geary representation that takes the minimum consumption level (7) explicitly into account, i.e. $u(c(t)) = \left( \left( c(t) - c^{\min}(t) \right)^{1-\sigma} - 1 \right) / (1-\sigma)$. Optimal consumption is then given by

$$c(\tau) = \frac{\rho - (1-\sigma) r}{\sigma} \left( a(\tau) + \frac{(1-\xi) \tilde{w}(\tau)}{r-g} \right) + \xi \tilde{w}(\tau)$$

where $\tilde{w}(\tau)$ is the wage from (28). Wealth therefore evolves according to

$$a(\tau) = \left( a(t) + \frac{(1-\xi) \tilde{w}(t)}{r-g} \right) e^{\gamma[\tau-t]} - \frac{(1-\xi) \tilde{w}(t)}{r-g} e^{g[\tau-t]}. \tag{D.10}$$

Hence, wealth $a(\tau)$ at some future point $\tau$ in time is monotonic in $a(t)$ in $t$.

Turning to the distribution of wealth in the future, the probabilities are linked again (see (25) or app. D.2.2) via $G(a(\tau)) = P(a(t))$, where $G(.)$ is the distribution function of $a(\tau)$ and $P(a(t))$ is the distribution function for $a(t)$. Densities are therefore linked by

$$g(a(\tau)) = \frac{d}{da(\tau)} P(a(t)) = p(a(t)) \frac{da(t)}{da(\tau)}.$$

Computing

$$a(t) = \left( a(\tau) + \frac{(1-\xi) \tilde{w}(t)}{r-g} e^{g[\tau-t]} \right) e^{-\frac{r-\rho}{\sigma}[\tau-t]} - \frac{(1-\xi) \tilde{w}(t)}{r-g}$$

shows that

$$\frac{da(t)}{da(\tau)} = e^{-\frac{r-\rho}{\sigma}[\tau-t]}.$$

The density $g(a(\tau))$ of wealth at $\tau$ is therefore described by

$$\begin{aligned} g(a(\tau)) &= p(a(t)) e^{-\frac{r-\rho}{\sigma}[\tau-t]} \\ &= p\left( \left( a(\tau) + \frac{(1-\xi) \tilde{w}(t)}{r-g} e^{g[\tau-t]} \right) e^{-\frac{r-\rho}{\sigma}[\tau-t]} - \frac{(1-\xi) \tilde{w}(t)}{r-g} \right) e^{-\frac{r-\rho}{\sigma}[\tau-t]}. \end{aligned} \tag{D.11}$$

- The steps in detail

(i) We start with the ex-ante-only-heterogeneity step. Starting from one initial density for wealth, $p(\hat{a}, t)$, obtained from (22), we solve for two densities $g(a(\tau))$ from (D.11), one for $r^{\text{low}}$ and for $r^{\text{high}}$. The wage is given by $\tilde{w}(t)$ as defined in (28).[54] We pin down $p_0$ by maximizing the

[54] We could also solve for 4 densities from 2 interest rates times two labour income levels, $w$ and $b$. As we are interested in *pure* capital income risk, we merge $w$ and $b$ to its mean. Otherwise we would allow for ex-ante heterogeneity in labour income and we would not be able to talk about pure capital income risk.

fit of a mixture of the two densities in fig. 6. In terms of (27), we maximize $F(t)$ by choosing $p_0$ in $g^{\text{model}}(a,t) = p_0 g_{\text{low}}(a,t) + (1-p_0) g_{\text{high}}(a,t)$. The fit is 64.7% only. The density is visible in fig. 12 below.

(ii) We now add ex-post heterogeneity. We compute densities from (D.11) for two interest rate *paths*. The paths start with the initial interest rate and switch to the other interest rate after 132 months (i.e. the half of 22 years). This gives an ex-ante-ex-post heterogeneity fit of 62.0%.[55] It falls compared to the ex-ante specification as ex-post uncertainty makes the resulting densities more equal. This is visible in fig. 12 below.

(iii) This step adds financial types: We solve (D.11) with the same interest rate paths used in the baseline model. This works as follows: (a) Starting again from the initial wealth distribution $p(\hat{a},t)$ and given one interest rate path, we solve for $g(a(\tau))$ from (D.11) employing the initial interest rate, at the $T$ where the interest rate jumps on this interest rate path. This gives $p(\hat{a},T)$. We then take this density and solve, employing the interest rate after the jump, for the remaining length of time $(22-T)$, where the total number of years is 22, to get the final $p(\hat{a},22)$. This yields one out of the $2n$ densities illustrated in fig. 5. (b) We repeat this for all paths. (c) We then choose the shares $p_j$ optimally. (The number of types $n$ is held constant.)

We obtain a fit of 65.9%. The difference between the overall fit and this fit provides the contribution of idiosyncratic risk in labour income.

- The findings

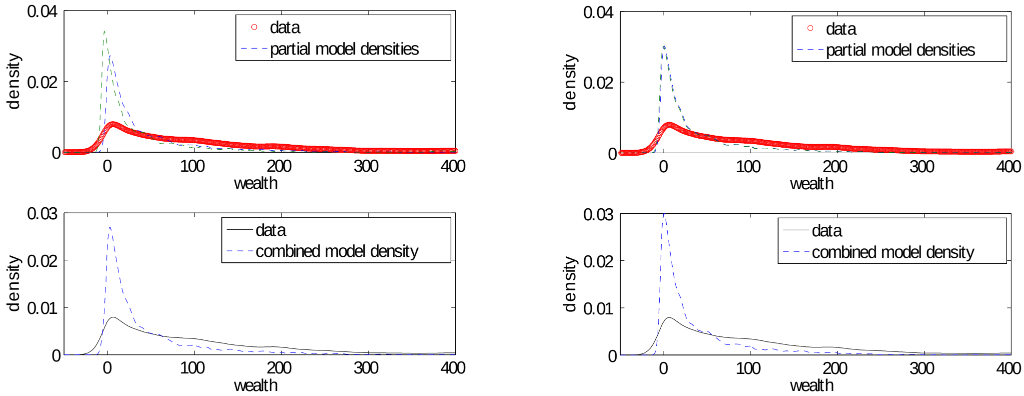


**Figure 12** *The density of wealth for pure capital income risk (at invariant labour income $\tilde{w}$): ex-ante heterogeneity (left) and ex-ante and ex-post heterogeneity with two interest rate paths (right)*

It seems surprising that the interest rate hardly has any impact on the density of wealth as visible in the left part of fig. 12. To check the plausibility of this finding, compute the ratio of wealth for two different interest rates based on (D.10). This ratio is

$$\frac{a(\tau, r_{\text{high}})}{a(\tau, r_{\text{low}})} = \frac{\left(a(t) + \frac{(1-\xi)\tilde{w}(t)}{r_{\text{high}}-g}\right) e^{\frac{r_{\text{high}}-\rho}{\sigma}[\tau-t]} - \frac{(1-\xi)\tilde{w}(t)}{r_{\text{high}}-g} e^{g[\tau-t]}}{\left(a(t) + \frac{(1-\xi)\tilde{w}(t)}{r_{\text{low}}-g}\right) e^{\frac{r_{\text{low}}-\rho}{\sigma}[\tau-t]} - \frac{(1-\xi)\tilde{w}(t)}{r_{\text{low}}-g} e^{g[\tau-t]}}. \tag{D.12}$$

For $\tau - t = 22$ years and $a(t) = 26,151$ $US\$$ (which is mean wealth in 1986), we get $a(\tau, r_{\text{high}})/a(\tau, r_{\text{low}}) = 1.43$. If we had a fixed initial wealth level, then the difference would be 43% after 22 years. This is roughly the value we get from looking at the left top figure in fig. 12.

[55] This fit depends of course on which specific financial type we look at. We chose the "representative" financial type, i.e. whose arrival rate is in the middle of all arrival rates we employ for the full quantitative solution.

When we look at $z = b$ and $z = w$ (and not $z = \tilde{w}$), we find the following.

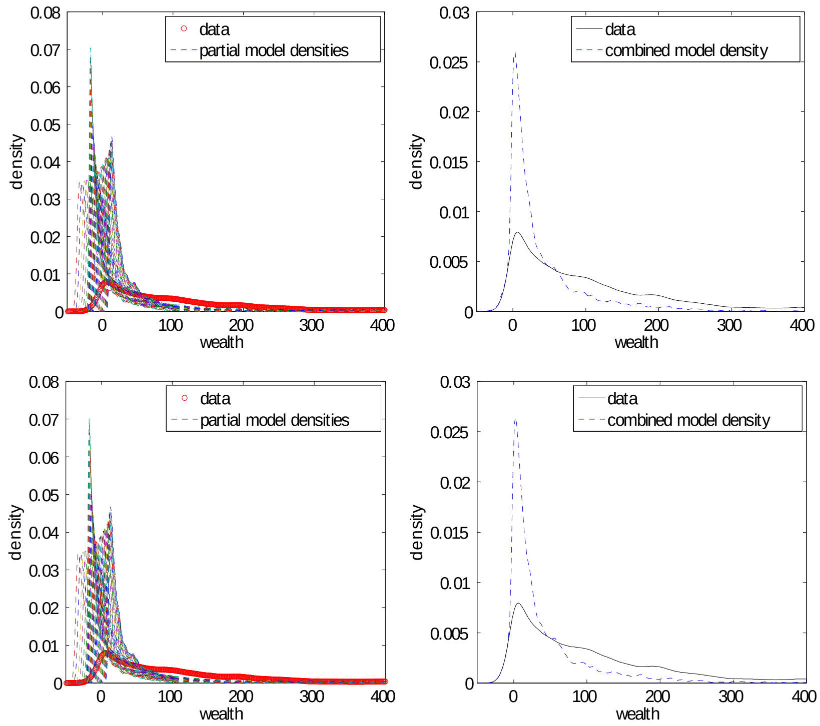


**Figure 13** *The effect of labour income on the density of wealth*

Here as well, just as with different interest rates in the ex-ante case, the effect of the level of labour income with all three channels of pure capital income risk seems to be negligible.

When we do a similar plausibility check as in (D.12) based on (D.10), we get

$$\frac{a(\tau, w)}{a(\tau, b)} = \frac{\left(a(t) + \frac{(1-\xi)w(t)}{r-g}\right) e^{\gamma[\tau-t]} - \frac{(1-\xi)w(t)}{r-g} e^{g[\tau-t]}}{\left(a(t) + \frac{(1-\xi)b(t)}{r-g}\right) e^{\gamma[\tau-t]} - \frac{(1-\xi)b(t)}{r-g} e^{g[\tau-t]}.} = \left\{ \begin{array}{c} 1.002 \\ 0.904 \end{array} \right\} \text{for } r = \left\{ \begin{array}{c} 4.5\% \\ 3.5\% \end{array} \right. \tag{D.13}$$

for the same parameter values as used above. The changes are therefore very small such that they are basically not visible in fig. 13.

### D.3.2 Baseline model without type heterogeneity

The importance of types is limited in these analyses with a small increase in the fit when we go from pure capital income risk (ex-ante, ex-post) to pure capital income risk with types (ex-ante, ex-post and types). What is the effect of removing types (i.e. having one type only) in the baseline model (where labour income risk is present as well)? The result is visible in fig. 14.

The fit for the baseline model with ex-ante and ex-post capital risk (but only one financial type) is in the best of all cases (where we pick the type that yields the highest fit) given by $F\,(2008) = 67.8\%$. The corresponding densities are shown in fig. 14. The worst fit with one type is 8.5%. The highest-fit type experiences 5.7 years in the high-interest-rate regime and 16.3 years in the low regime.

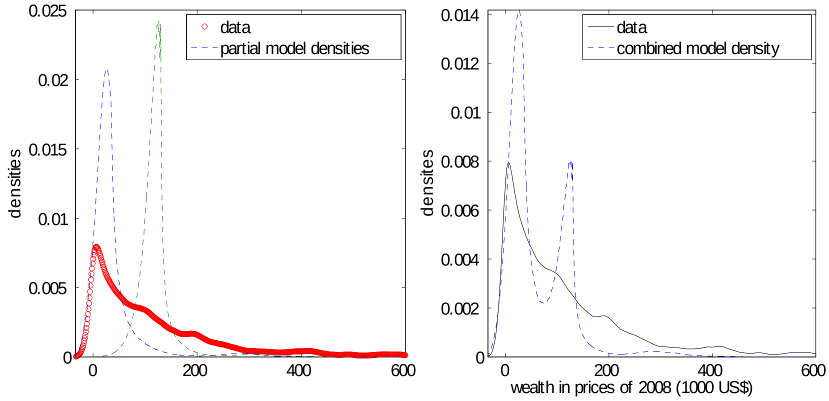


**Figure 14** *The density of wealth for two interest rate paths, i.e. one financial type*

### D.3.3 Pure capital income risk for a flexible interest rate distribution

We now ask how the fit changes when we generalize our two-point interest rate distribution. We now let the interest rates range from 4% to 15% with many realizations between 4% and 15%. We study two cases: ex-ante heterogeneity only and all three sources of capital income risk (ex-ante, ex-post and types).

- Ex-ante heterogeneity only

When we only allow for ex-ante heterogeneity, we fix $n$ interest rates between 4% and 15%. We compute one partial density per interest rate according to (D.11). We then choose weights $p_i$ such that the fit is maximized. This gives a certain fit. We then choose $n$ such that the fit is maximized. The result is in the next figure.

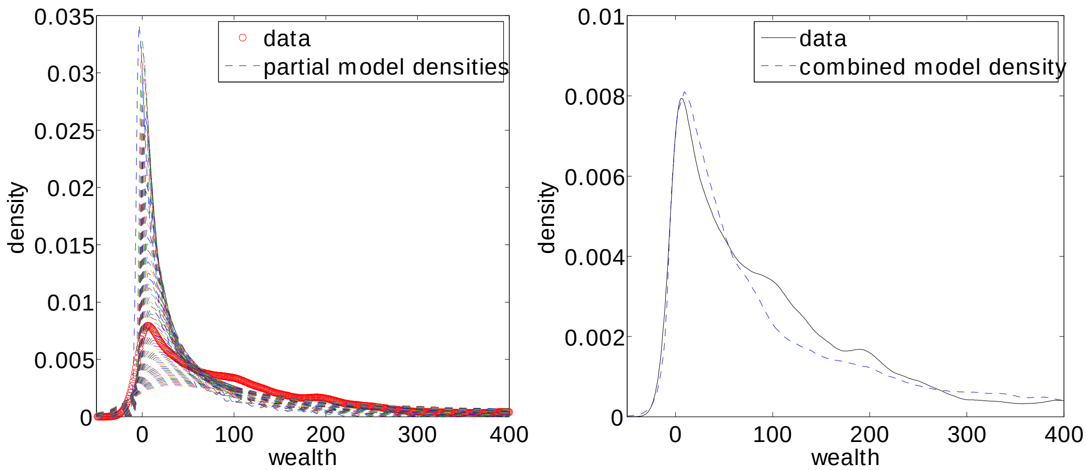


**Figure 15** *The fit for ex-ante heterogeneity in constant interest rates*

The optimal number of paths is 69 (out of 200), and the fit is 89.8%. The link between $n$ and the fit is shown in the next figure.

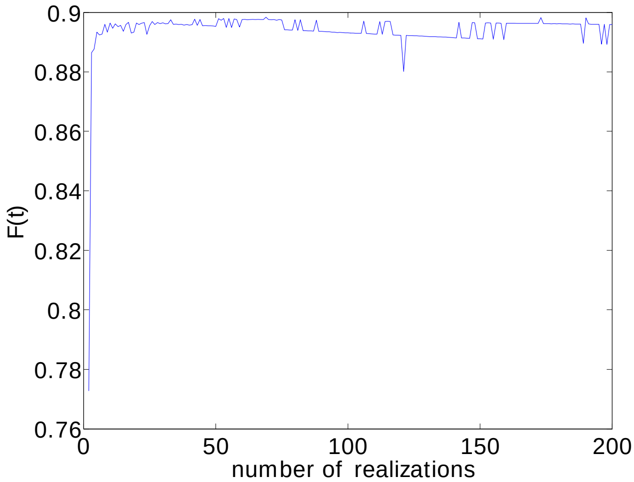


**Figure 16** *The fit $F(t)$ as a function of the number of paths*

- Three sources of risk

When we allow for all sources of pure capital income risk (ex ante, ex post, types), we take the $n$ interest rate paths from the baseline model. We compute a first density for the end of the first subperiod of one path according to (D.11). The resulting density is the initial density for the second subperiod. By doing so, we end up with $n$ densities after 22 years. We compute optimal weights $p_j$. We can eventually match the 2008 density by 96.7% as shown in fig. 17. It seems that labour income uncertainty is not required.

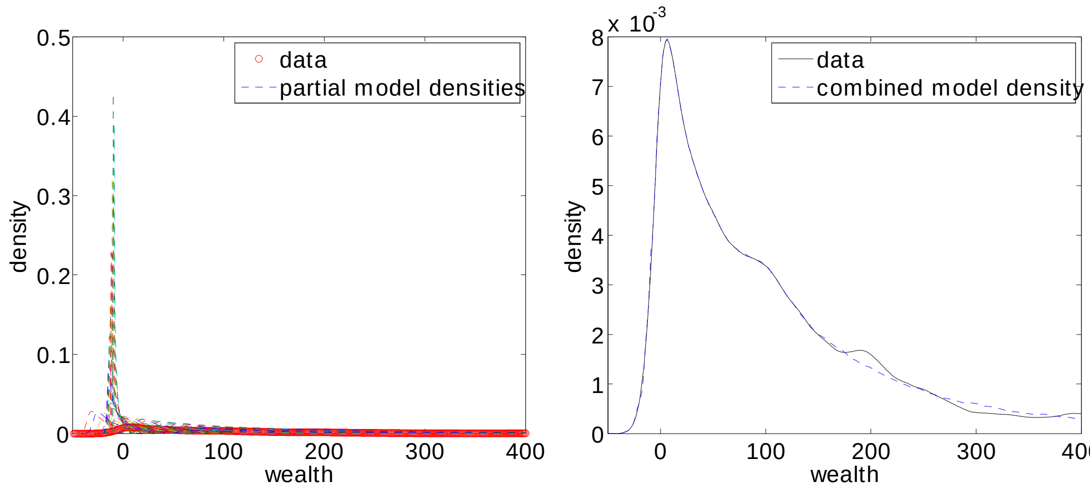


**Figure 17** *The optimal fit with pure capital risk - do we need labour income risk at all?*

### D.3.4 High interest rate $r^{\mathbf{high}} = 8\%$

We report here the findings of our robustness check.

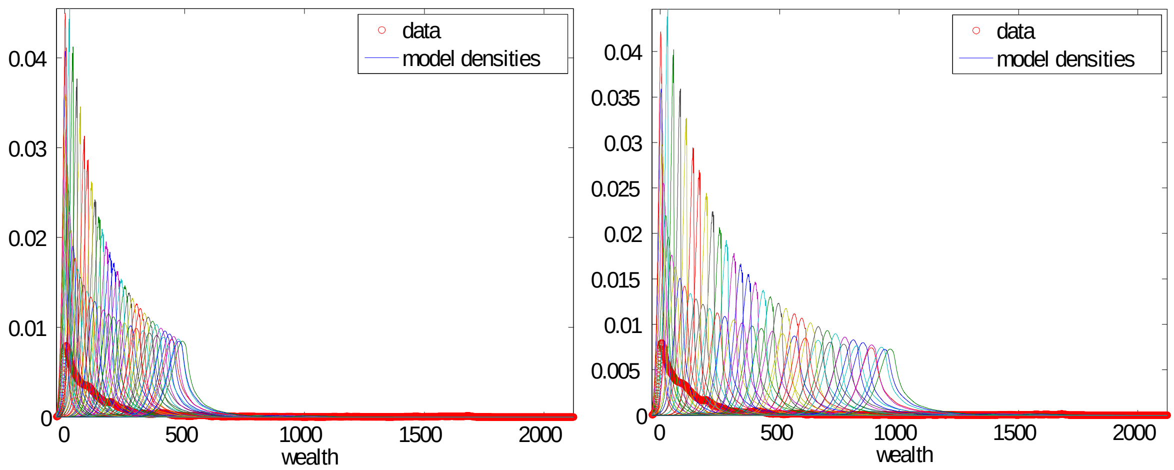


**Figure 18** *The effect on wealth distributions by financial types for low and high initial interest rates ($r^{high} = 4.5\%$ in left figure and $r^{high} = 8\%$ in right figure)*

The figure shows the densities for the interest rate paths after 22 years (as in fig. 5 in the main text). The higher interest rate clearly implies that densities move further to the right.

### D.3.5 Varying risk aversion

How does the fit look like when $\sigma = 0.8$ or $\sigma = 1.2$?

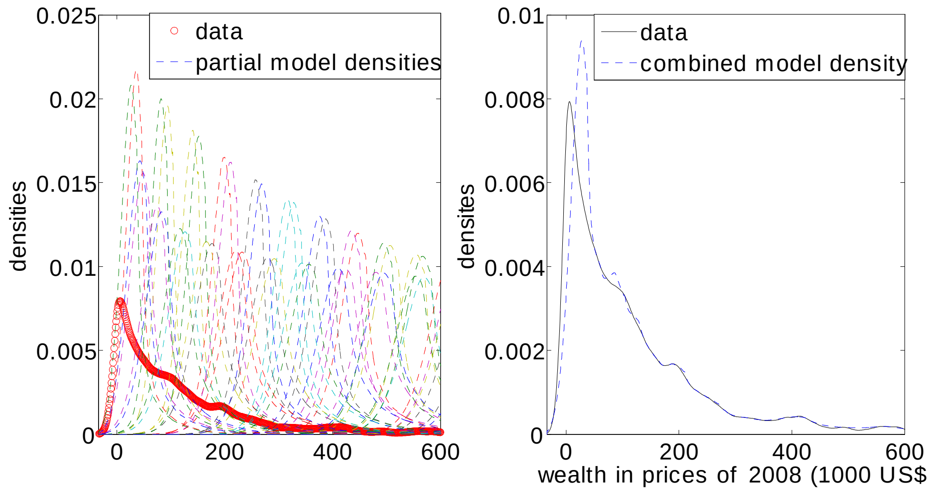


**Figure 19** *Partial densities and overall fit for $\sigma = 0.8$*

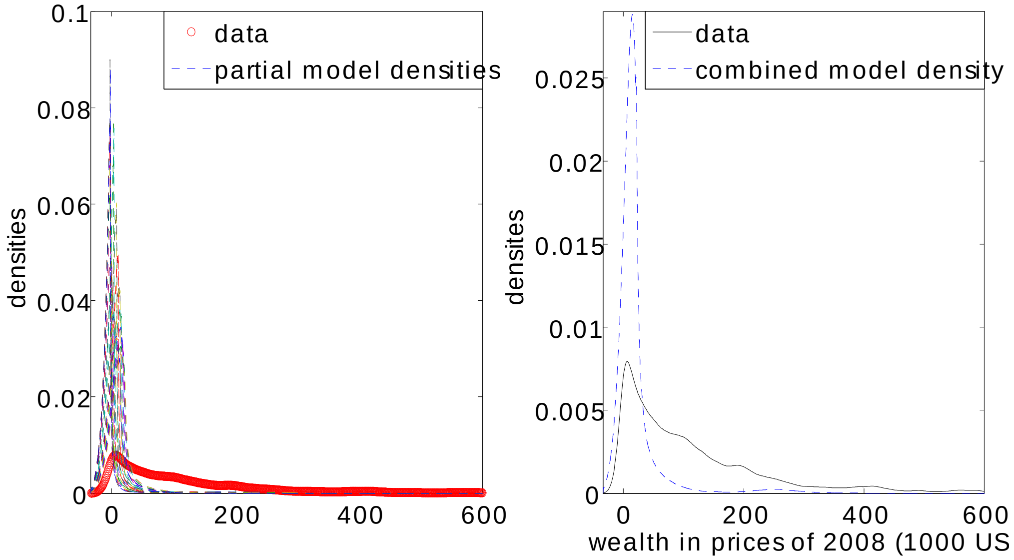


**Figure 20** *Partial densities and overall fit for* $\sigma = 1.2$ *(Note that there is no exploding regime for this value of* $\sigma = 1.2$*)*

When we look at the corresponding quantitative phase diagrams – in analogy to fig. 4 in the main text – we can easily understand the effect of risk aversion.

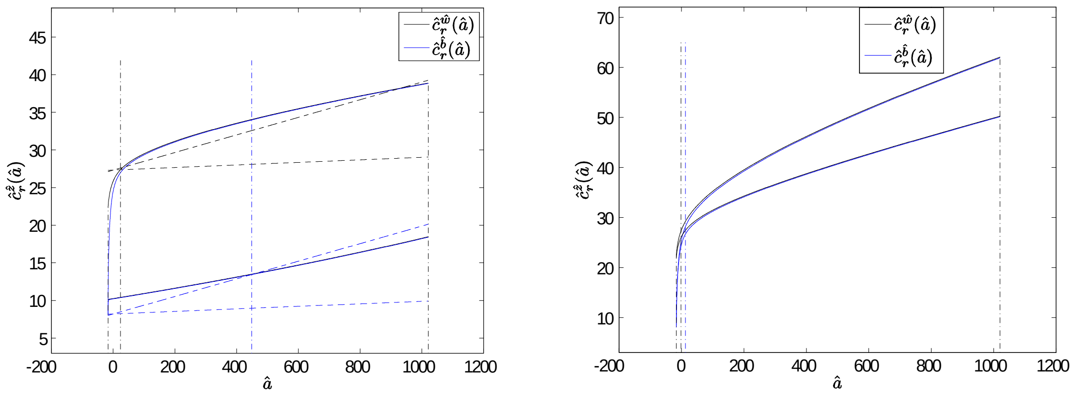


**Figure 21** *Consumption paths for* $\sigma = 0.8$ *(left) and* $\sigma = 1.2$ *(right)*

For a low $\sigma$ of 0.8, we have a standard regime and an exploding regime. The latter has too low consumption levels, however, such that wealth densities move to the right too quickly. As a consequence, individuals are "too rich" and the combined model density in fig. 19 is too far to the right. By contrast, for $\sigma = 1.2$, there is no exploding regime visible in the right quantitative phase diagram of fig. 21. As a consequence, all wealth densities in fig. 20 are concentrated around $a_w^*$ and the combined model density hardly has any probability mass in the right tail.

### D.3.6 Wealth percentiles

We present wealth shares in the data, $\omega^{\text{data}}_{\text{percentile},\, t}$, and model, $\omega^{\text{model}}_{\text{percentile},\, t}$, for selected percentiles of the population. Note that these wealth shares are not targeted but are the result of targeting either the wealth density in 2008 or the average over densities in all years.

- Findings for target year 2008

| percentile | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 95 | 99 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| data | −0.3 | 0.3 | 1.9 | 4.7 | 8.9 | 14.9 | 22.9 | 34.1 | 50.8 | 64.9 | 89.7 |
| model | −0.6 | 0.1 | 2.1 | 5.6 | 11.1 | 18.8 | 29.2 | 43.3 | 63.2 | 78.0 | 94.1 |

**Table 6** *Wealth shares (in %) for 2008 in the data and model when targeting the density in year 2008*

The average of the deviations for the wealth share in 2008,

$$\Delta_{2008} \equiv \Sigma_{\text{percentile}=1}^{99} \left(\omega_{\text{percentile, 2008}}^{\text{data}} - \omega_{\text{percentile, 2008}}^{\text{model}}\right) /99, \tag{D.14}$$

is −4.0%. The average for wealth share fit over all waves,

$$\Delta_{\text{all waves}} \equiv \Sigma_{\text{waves}=1}^{12} \Sigma_{\text{percentile}=1}^{99} \left(\omega_{\text{percentile, }t}^{\text{data}} - \omega_{\text{percentile, }t}^{\text{model}}\right) /99/12, \tag{D.15}$$

is −7.6%.

- Findings when targeting all years

| percentile | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 95 | 99 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| data | −0.3 | 0.3 | 1.9 | 4.7 | 8.9 | 14.9 | 22.9 | 34.1 | 50.8 | 64.9 | 89.7 |
| model | −0.4 | 0.2 | 1.8 | 5.4 | 11.5 | 20.3 | 32.6 | 49.1 | 70.4 | 83.2 | 95.4 |

**Table 7** *Wealth shares (in %) for 2008 in the data and model when targeting all years*

When targeting all years, the average of the wealth share fit for 2008 from (D.14) is −5.7%. The average of the wealth share fit over all waves from (D.15) is −2.6%.

### D.3.7 Is financial ability time-invariant?

The figures show in more detail why up to 1998 average time spent in the high regime is higher than up to 2008.

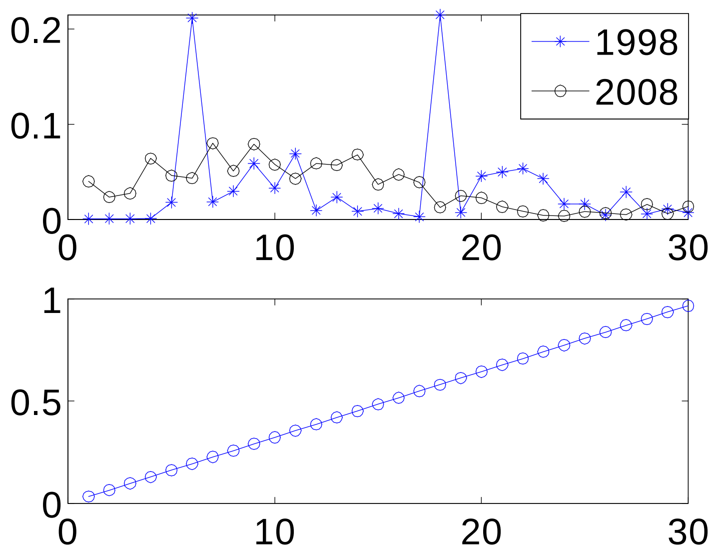


**Figure 22** *The share $p_i$ of financial types $i$ (horizontal axis) in the upper figure and the share of time that the corresponding types $i$ are in the high-interest rate regime in the lower figure*

The text reports $\text{share}^{1998} \equiv \Sigma_{j=1}^{26} p_j^{1998} \text{share}_j$ and a corresponding $\text{share}^{2008}$ for 2008.

## D.4 The distribution of idiosyncratic interest rates

The mean $\mu$ and variance $\sigma^2$ for a discrete random variable $X$ are by definition given by $\mu = \Sigma_{i=1}^{n} \pi_i x_i$ and $\sigma^2 = \Sigma_{i=1}^{n} \pi_i [x_i - \mu]^2$ . Hence, the numerical mean and standard deviation in month $m$ *for one financial type* are computed as

$$\begin{aligned} \hat{\mu}_{r,m} &= \hat{\pi}_m r_{\text{high}} + (1 - \hat{\pi}_m)\, r_{\text{low}}, \qquad m = 1...264, \\ \hat{\sigma}_{r,m} &= \sqrt{\hat{\pi}_m [r_{\text{high}} - \mu_{r,m}]^2 + (1 - \hat{\pi}_m)[r_{\text{low}} - \mu_{r,m}]^2} \end{aligned} \tag{D.16}$$

where $\hat{\pi}_m$ in the code is computed as $\hat{\pi}_m = \Sigma_{j=1}^{2n} p_j I_{j,m}$ and $I_{j,m} = 1$ when path $j$ in month $m$ has a high return and 0 otherwise. We thereby obtain 264 means and standard deviations. The last moments corresponding to the target year (i.e. December) are reported. For example, when 2008 is targeted, the last elements of the vector and their mean are reported.

# E Numerical implementation

## E.1 The evolution of densities of wealth and interest rate jumps

As individuals are myopic with respect to interest rate changes (see footnote 23), we have "only" a two-dimensional FPE system in (23). We describe here how changes in the interest rate are implemented numerically.

- The implementation of interest rate jumps

We assume that the individual stays in the low-interest-rate regime and randomly jumps to the high-interest-rate regime. If the individual stays 24 months in the low-interest-rate regime, then they spend 240 (= 264 − 24) months in the high-interest-rate regime. The possible durations that an individual can stay in an interest rate regime are from 2 months to 262 months.

In principle, jumps in interest rates follow financial types. In order to facilitate the numerical implementation, we translate duration rates (e.g. 1/240) into expected duration (e.g. 240 months). The link between the arrival rates as described after (2) is implemented by

$$\lambda^{\text{high}} = \frac{1}{264 - 1/\lambda^{\text{low}}}.$$

We then solve for the distribution of wealth after the initial random duration in the low-interest-rate regime, then for the random duration in the high-interest-rate regime. Then, 22 years are over.

- Combining densities

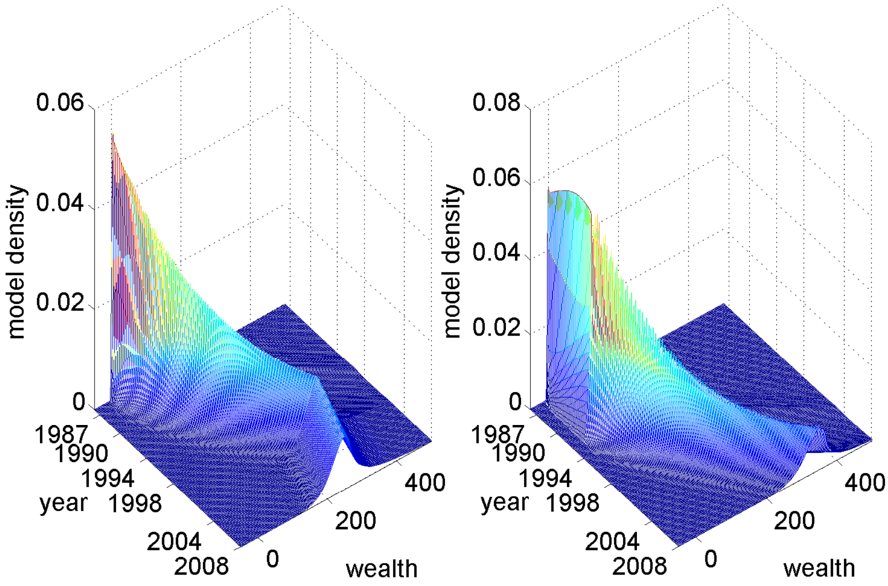


**Figure 23** *The evolution of the density of wealth when the interest rate jumps once*

The left figure shows an individual starting with a high interest rate. The interest rate drops at 2000 to the low level. In the right figure, the individual starts with a low interest rate and the interest rate jumps upwards in 1994.

## E.2 Solving Fokker-Planck equations by the method of characteristics

The Fokker-Planck equations (FPEs) associated with our individual's maximization problem read

$$\frac{\partial}{\partial t}p^{\hat{w}}(\hat{a},t)+\left[(r-g)\hat{a}+\hat{w}-\hat{c}^{\hat{w}}(\hat{a})\right]\frac{\partial}{\partial\hat{a}}p^{\hat{w}}(\hat{a},t)=\left[\frac{dc^{\hat{w}}(\hat{a})}{d\hat{a}}-(r-g)-s\right]p^{\hat{w}}(\hat{a},t)+\mu p^{\hat{b}}(\hat{a},t), \tag{E.1}$$

$$\frac{\partial}{\partial t}p^{\hat{b}}(\hat{a},t)+\left[(r-g)\hat{a}+\hat{b}-\hat{c}^{\hat{b}}(\hat{a})\right]\frac{\partial}{\partial\hat{a}}p^{\hat{b}}(\hat{a},t)=sp^{\hat{w}}(\hat{a},t)+\left[\frac{dc^{\hat{b}}(\hat{a})}{d\hat{a}}-(r-g)-\mu\right]p^{\hat{b}}(\hat{a},t). \tag{E.2}$$

We solve these FPEs using the method of characteristics building on earlier work by Nagel (2013, ch. 5). Solving a FPE involves locating a curve along which the solution for the densities follows an ordinary differential equation (ODE). Such a curve is called "characteristic curve" or simply a "characteristic". These ODEs can be solved on a rectangle $[\hat{a}^{\text{nat}},\hat{a}^{\max}]\times[0,T]$ using the empirically given initial conditions $p^{\hat{w}}(\hat{a},0)$ and $p^{\hat{b}}(\hat{a},0)$ and boundary conditions. The latter are given, in our case, by $p^{\hat{w}}(\hat{a}^{\text{nat}},t)=0$ and by $p^{\hat{b}}(\hat{a}^{\max},t)=0$.[56]

[56] Such a PDE system with initial and boundary conditions is called "initial-boundary value problem". See Strikwerda (2004) for a general background on PDEs, on understanding whether such systems are well-posed and especially on finite difference schemes for numerically solving PDEs.

A detailed discussion of the method of characteristics in the context of partial differential equations can be found in Mattheij et al. (2005, ch. 2.2). The characteristic equations associated with the FPEs above are given by (see Nagel, 2013, ch. 5)

$$\frac{d\hat{a}}{dt} = (r-g)\hat{a} + \hat{w} - \hat{c}^{\hat{w}}(\hat{a}), \tag{E.3}$$

$$\frac{d\hat{a}}{dt} = (r-g)\hat{a} + \hat{b} - \hat{c}^{\hat{b}}(\hat{a}), \tag{E.4}$$

$$\frac{dp^{\hat{w}}(\hat{a},t)}{dt} = \left[\frac{dc^{\hat{w}}(\hat{a})}{d\hat{a}} - (r-g) - s\right] p^{\hat{w}}(\hat{a},t) + \mu p^{\hat{b}}(\hat{a},t), \tag{E.5}$$

$$\frac{dp^{\hat{b}}(\hat{a},t)}{dt} = \left[\frac{dc^{\hat{b}}(\hat{a})}{d\hat{a}} - (r-g) - \mu\right] p^{\hat{b}}(\hat{a},t) + sp^{\hat{w}}(\hat{a},t). \tag{E.6}$$

The solutions to $p^{\hat{w}}(\hat{a},t)$ and $p^{\hat{b}}(\hat{a},t)$ obtained from (E.5) and (E.6) hold along the characteristic curves (E.3) and (E.4), respectively. Detailed discussion on the numerical method solving these characteristic equations can also be found in Nagel (2013, ch. 5).

An issue with modelling densities and their numerical solution is the question of mass points. In fact, our dynamic system $(\hat{a},\hat{z})$ from (11) and (12) could imply a masspoint at the lower bound $\hat{a}^{\text{nat}}$ of wealth if consumption $\hat{c}_r^{\hat{b}}(\hat{a}(t))$ of the unemployed reaches $\hat{a}^{\text{nat}}$ in finite time. We do not model mass points explicitly as empirical densities tend to be smooth and as we found them negligible from a quantitative perspective in our calibrations.[57] Mass points would not result for an endogenous exit rate from unemployment (as in Lise, 2013) where getting closer to $\hat{a}^{\text{nat}}$ increases the exit rate (which goes to infinity as wealth approaches $\hat{a}^{\text{nat}}$).

[57] See Birkner and Wälde (2014) for FPE systems that include additional ODEs that describe mass points. See Achdou et al (2017) for a discussion of mass points in a continuous time Bewley-Huggett-Aiyagari model.