Online Appendix

Stock Price Booms and Expected Capital Gains

Klaus Adam, Albert Marcet and Johannes Beutel

APPENDIX

A1. Data Sources

Stock price data: our stock price data is for the United States and has been downloaded from 'The Global Financial Database' (http://www.globalfinancialdata.com). The period covered is Q1:1946-Q1:2012. The nominal stock price series is the 'SP 500 Composite Price Index (w/GFD extension)' (Global Fin code '_SPXD'). The daily series has been transformed into quarterly data by taking the index value of the last day of the considered quarter. To obtain real values, nominal variables have been deflated using the 'USA BLS Consumer Price Index' (Global Fin code 'CPUSAM'). The monthly price series has been transformed into a quarterly series by taking the index value of the last month of the considered quarter. Nominal dividends have been computed as follows

$$D_t = \left(\frac{I^D(t)/I^D(t-1)}{I^{ND}(t)/I^{ND}(t-1)} - 1 \right) I^{ND}(t)$$

where I^{ND} denotes the 'SP 500 Composite Price Index (w/GFD extension)' described above and I^D is the 'SP 500 Total Return Index (w/GFD extension)' (Global Fin code '_SPXTRD '). We first computed monthly dividends and then quarterly dividends by adding up the monthly series. Following Campbell 2003, dividends have been deseasonalized by taking averages of the actual dividend payments over the current and preceding three quarters.

Interest rate data: As nominal interest rate we use the 90 Days T-Bills Secondary Market (Global Fin code ITUSA3SD). The weekly (to the end of 1953) and daily (after 1953) series has been transformed into a quarterly series using the interest rate corresponding to the last week or day of the considered quarter and is expressed in quarterly rates (not annualized). To obtain real values, nominal variables have been deflated using the 'USA BLS Consumer Price Index' (Global Fin code 'CPUSAM').

Stock market survey data: The UBS survey is the UBS Index of Investor Optimism, which is available (against a fee) at

http://www.ropercenter.uconn.edu/ data_access/data/datasets/ubs_investor.html.

The quantitative question on stock market expectations has been surveyed over the period Q2:1998-Q4:2007 with 702 responses per month on average and has thereafter been suspended. For each quarter we have data from three monthly surveys, except for the first four quarters and the last quarter of the survey period where we have only one monthly survey per quarter. The Shiller survey data covers individual investors over the period Q1:1999Q1-Q4:2012 and has been kindly made available to us by Robert Shiller at Yale University. On average 73 responses per quarter have been recorded for the question on stock price growth. Since the Shiller data refers to the Dow Jones, we used the PD ratio for the Dow Jones, which is available at http://www.djaverages.com/, to compute correlations. The CFO survey is collected by Duke University and CFO magazine and collects responses from U.S. based CFOs over the period Q3:2000-Q4:2012 with on average 390 responses per quarter, available at http://www.cfosurvey.org/ .

Inflation expectations data: The Survey of Professional Forecasters (SPF) is available from the Federal Reserve Bank of Philadelphia at http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/. The Michigan Surveys of Consumers are collected by Thomson Reuters/University of Michigan

(http://www.sca.isr.umich.edu/).

A2. Correlations between PD Ratio and Actual/Survey Returns

This appendix documents for the UBS Survey, the CFO survey and the Shiller individual investor survey that there exists a positive correlation between the PD ratio and survey expected returns (or capital gains), but a negative correlation between the PD ratio and actual returns (or capital gains).

Table A1 documents the positive correlation for the UBS survey. Results are independent of how one extracts expectations from the survey (using the median or the mean expectation, using inflation expectations from the Michigan survey or the Survey of Professional Forecasters (SPF) to obtain real return expectations, using plain nominal returns instead or real returns, or when restricting attention to investors with more than 100.000 US$ in financial wealth). The numbers reported in brackets in table A1 (and in subsequent tables) are autocorrelation robust p-values for the hypothesis that the correlation is smaller or equal to zero.[98] These p-values are *not adjusted for small sample bias*, as there is no generally accepted approach for how to perform such adjustments. This said, the p-values for the null hypothesis are all below the 5% significance level and in many cases below the 1% level.

A positive correlation is equally obtained when considering other survey data. Table A2 reports the correlations between the PD ratio and the stock price growth expectations from Bob Shiller's Individual Investors' Survey.[99] The table shows that price growth expectations are also strongly positively correlated with the PD ratio, suggesting that the variation in expected returns observed in the UBS survey is due to variations in expected capital gains. Table A2 also shows that correlations seem to become stronger for longer prediction horizons.

Table A3 reports the correlations for the stock return expectations reported in the Chief Financial Officer (CFO) survey which surveys chief financial officers from large U.S. corporations. Again, one finds a strong positive correlation.

Table A4 reports the correlations between the PD ratio and the realized real returns (or capital gains) in the data, using the same sample periods as are available for the surveys considered in tables A1 to A3, respectively. The point estimate for the correlation is negative in all cases, although the correlations fall short of being significant the 5% level due to the short sample length for which the survey data is available.

[98] The sampling width is four quarters, as is standard for quarterly data, and the test allows for contemporaneous correlation, as well as for cross-correlations at leads und lags. The p-values are computed using the result in Roy 1989.

[99] Shiller's price growth data refers to the Dow Jones Index. The table thus reports the correlation of the survey measure with the PD ratio of the Dow Jones.

UBS Gallup	**Nominal Return Exp.**		**Real Ret. Exp. (SPF)**		**Real Ret. Exp. (Michigan)**	
	Average	**Median**	**Average**	**Median**	**Average**	**Median**
Own portfolio, >100k US$	0.80 (0.01)	0.78 (0.01)	0.79 (0.01)	0.77 (0.01)	0.84 (0.01)	0.83 (0.01)
Own portfolio, all investors	0.80 (0.01)	0.76 (0.02)	0.79 (0.01)	0.75 (0.02)	0.84 (0.01)	0.80 (0.01)
Stock market, >100k US$	0.90 (0.03)	0.89 (0.04)	0.90 (0.03)	0.88 (0.03)	0.91 (0.03)	0.88 (0.03)
Stock market, all investors	0.90 (0.03)	0.87 (0.04)	0.90 (0.03)	0.87 (0.04)	0.91 (0.03)	0.88 (0.03)

Table A1: Correlation between PD ratio and 1-year ahead expected return measures
(UBS Gallup Survey, robust p-values in parentheses,
without small sample correction for p-values)

Shiller Survey	**Nominal Capital Gain Exp.**		**Real Capital Gain. Exp. (SPF)**		**Real Capital Gain Exp. (Michigan)**	
Horizon	**Average**	**Median**	**Average**	**Median**	**Average**	**Median**
1 month	0.46 (0.01)	0.48 (0.01)	0.45 (0.01)	0.47 (0.01)	0.46 (0.01)	0.49 (0.01)
3 months	0.57 (0.01)	0.64 (0.00)	0.54 (0.01)	0.61 (0.00)	0.56 (0.01)	0.62 (0.01)
6 months	0.58 (0.01)	0.75 (0.01)	0.54 (0.02)	0.70 (0.01)	0.56 (0.02)	0.71 (0.01)
1 year	0.43 (0.03)	0.69 (0.01)	0.38 (0.05)	0.62 (0.01)	0.42 (0.04)	0.64 (0.02)
10 years	0.74 (0.01)	0.75 (0.01)	0.66 (0.02)	0.71 (0.01)	0.71 (0.02)	0.75 (0.01)

Table A2: Correlation between PD ratio and expected stock price growth
(Shiller's Individual Investors' Survey, robust p-values in parentheses,
without small sample correction for p-values)

CFO Survey	**Nominal Return Exp.**		**Real Return Exp. (SPF)**		**Real Return Exp. (Michigan)**	
	Average	**Median**	**Average**	**Median**	**Average**	**Median**
1 year	0.71 (0.00)	0.75 (0.00)	0.62 (0.00)	0.69 (0.00)	0.67 (0.00)	0.72 (0.00)

Table A3: Correlation between PD ratio and 1-year ahead expected stock return measures
(CFO Survey, robust p-values in parentheses, without small sample correction for p-values)

Variables	Time Period	Stock Index	Correlation
PD, 1 year-ahead real return	UBS Gallup sample (stock market exp.)	S&P 500	-0.66 (0.08)
PD, 1 year-ahead real price growth	Shiller 1 year sample	Dow Jones	-0.42 (0.06)
PD, 10 year-ahead real price growth	Shiller 10 year sample	Dow Jones	-0.88 (0.16)
PD, 1 year-ahead real return	CFO sample	S&P 500	-0.46 (0.06)

Table A4: Correlation between PD and actual real returns/capital gains
(robust p-value in parentheses, without small sample correction for p-values)

A3. Proof of Proposition 1

Proof of part a)
Under the null hypothesis of rational expectations ($E_t^{\mathcal{P}} = E_t$) equation (1) implies

$$R_{t,t+N} = a^N + c^N \frac{P_t}{D_t} + u_t^N + \varepsilon_t^N, \tag{A1}$$

where ε_t^N is the prediction error $R_{t,t+N} - E_t R_{t,t+N}$ from the *true* data-generating process, the conditional expectation is taken with respect to investors' information at t. Since P_t/D_t is in this information set under RE and given (2) we have

$$E\left[x_t\left(u_t^N + \varepsilon_t^N\right)\right] = 0. \tag{A2}$$

Therefore, $u_t^N + \varepsilon_t^N = \mathbf{u}_t^N$ and the null hypothesis of rational expectations implies

$$c^N = \mathbf{c}^N. \tag{A3}$$

Equations (3) (5) define a SUR system of equations with dependent variables $\mathcal{E}_t^N$ and $R_{t,t+N}$, and explanatory variables in both equations $x_t = (1, \frac{P_t}{D_t})'$. Under the null hypothesis the error terms satisfy the orthogonality conditions (4) and (A2).

For part *a*) of Proposition 1 we define the OLS estimator equation by equation $\widehat{\beta}_T$ as

$$\widehat{\beta}_T \equiv \begin{bmatrix} \widehat{a}_T^N \\ \widehat{c}_T^N \\ \widehat{\mathbf{a}}_T^N \\ \widehat{\mathbf{c}}_T^N \end{bmatrix} = \left(I_2 \otimes \sum_{t=1}^{T} x_t x_t' \right)^{-1} \sum_{t=1}^{T} \begin{bmatrix} \mathcal{E}_t^N \\ R_{t,t+N} \end{bmatrix} \otimes x_t,$$

where I_2 is a 2×2 identity matrix. A standard result ensures that under the assumptions OLS equation by equation is consistent and efficient among the set of estimators that use orthogonality conditions (4) and (A2).

As is well known, as $T \to \infty$ we have

$$\text{(A4)} \qquad \sqrt{T}\left(\widehat{\beta}_T - \beta_0\right) \to N\left(0,\ \left[I_2 \otimes E(x_t x_t')\right]^{-1}\ S_w\ \left[I_2 \otimes E(x_t x_t')\right]^{-1}\right),$$

in distribution, where

$$\begin{aligned} S_w &= \Gamma_0 + \sum_{k=1}^{\infty}\left(\Gamma_k + \Gamma_k'\right) \\ \Gamma_k &= E\left(\begin{bmatrix} u\mu_t \\ u\varepsilon_t \end{bmatrix}\left[u\mu_{t-k}, u\varepsilon_{t-k}\right] \otimes x_t x_{t-k}'\right), \end{aligned}$$

where $u\mu_t \equiv u_t^N + \mu_t^N$ and $u\varepsilon_t \equiv u_t^N + \varepsilon_t^N$. The footnote of the proposition contains all boundedness conditions required to ensure validity of asymptotic distribution, $E(x_t x_t')$ is invertible because $var(P_t/D_t) > 0$.

To build the test-statistic, we only need to find an estimator for var-cov matrix in (A4). We estimate $E\left(x_t x_t'\right)$ by $\frac{1}{T}\sum_{t=1}^{T} x_t x_t'$. Since $u\mu_t$ and $u\varepsilon_t$ are not forecasting errors, there is no reason why Γ_k should be zero for any k, so we use a Newey-West estimator $\widehat{S}_w$. Therefore, postmultiplying $\left(\beta_T - \beta_0\right)$ by $[0, 1, 0, -1]'$ in (A4), and letting

$$\text{(A5)} \quad \widehat{\sigma}^2_{c-c} \equiv \frac{1}{T}[0,1,0,-1]\left[I_2 \otimes \frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right]^{-1} \widehat{S}_w \left[I_2 \otimes \frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right]^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

we have that under the null hypothesis

$$\sqrt{T}\,\frac{\widehat{c}_T^N - \widehat{\mathbf{c}}_T^N}{\widehat{\sigma}_{c-c}} \to N(0,1) \text{ in distribution.}$$

Proof of part b)

Equations (3) (5) in the current paper are each of the form of equation (1) in Stambaugh 1999. Focusing first on (3), our $(\mathcal{E}_t^N, u_t + \mu_t, PD_t, \varepsilon_{t+1}^{PD})$ play the role of (y_t, u_t, x_{t-1}, v_t) in Stambaugh. Note, in particular, that to match his framework we need to have PD_t play the role of x_{t-1}, implying that our ε_{t+1}^{PD} plays the role of Stambaugh's v_t. Therefore, assumption at the bottom of page 378 in Stambaugh requires that $(u_t^N + \mu_t^N, \varepsilon_{t+1}^{PD})$ is jointly normal. Under normality, using orthogonality of measurement error, it follows from proposition 4 in Stambaugh 1999 that

$$\text{(A6)} \qquad E(\widehat{c}_T^N - c^N) = \frac{cov(\varepsilon_{t+1}^{PD}, u_t^N)}{var(\varepsilon_t^{PD})} E(\widehat{\rho}_T - \rho)$$

where $\widehat{\rho}_T$ is the OLS estimator of ρ. Since u_t contains information that under the null is useful for predicting future returns we could expect $\frac{cov(\varepsilon_{t+1}^{PD}, u_t^N)}{var(\varepsilon_t^{PD})} \neq 0$ and the bias to be non-zero. Similarly, we have,

$$E(\widehat{\mathbf{c}}_T^N - \mathbf{c}^N) = \frac{cov(\varepsilon_{t+1}^{PD}, u_t^N + \varepsilon_t^N)}{var(\varepsilon_t^{PD})} E(\widehat{\rho}_T - \rho) \tag{A7}$$

$$= E(\widehat{c}_T^N - c^N) + \frac{cov(\varepsilon_{t+1}^{PD}, \varepsilon_t^N)}{var(\varepsilon_t^{PD})} E(\widehat{\rho}_T - \rho). \tag{A8}$$

A4. Parameterization of the Wage Process

We set $1 + \rho$ equal to the average consumption-dividend ratio in the U.S. over the period 1946-2011, using the 'Personal Consumption Expenditures' and 'Net Corporate Dividends' series from the Bureau of Economic Analysis. This delivers $\rho = 22$. The consumption-dividend ratio fluctuates considerably over time and displays a close to unit root behavior, with the quarterly sample autocorrelation being equal to 0.99, prompting us to consider only values close to one for the persistence parameter p.

Following Campbell and Cochrane 1999, our remaining calibration targets are

$$\sigma_{c,t} = \frac{1}{7}\sigma_D \tag{A9}$$

and

$$\rho_{c,d,t} = 0.2, \tag{A10}$$

where $\sigma_{c,t}$ denotes the conditional standard deviation of log consumption growth and $\rho_{c,d,t}$ the conditional correlation between log consumption growth and log dividend growth. Aggregate consumption is given by $C_t = D_t + W_t$ so that

$$\begin{aligned}\log \frac{C_t}{C_{t-1}} &= \log \frac{D_t\left(1 + \frac{W_t}{D_t}\right)}{D_{t-1}\left(1 + \frac{W_{t-1}}{D_{t-1}}\right)} \\ &= c + \log \varepsilon_t^D + \log \varepsilon_t^W - (1-p)\log\left(1 + \frac{W_{t-1}}{D_{t-1}}\right),\end{aligned}$$

where c summarizes constant terms. Conditional variance of log consumption growth is thus equal to

$$\sigma_{c,t}^2 = \sigma_D^2 + \sigma_W^2 + 2\sigma_{DW} \tag{A11}$$

and conditional covariance between log consumption and dividend growth given by

$$\begin{aligned}\mathrm{cov}_{t-1}\left(\log \frac{C_t}{C_{t-1}}, \log \frac{D_t}{D_{t-1}}\right) &= \mathrm{cov}_{t-1}\left(\log \varepsilon_t^D + \log \varepsilon_t^W, \log \varepsilon_t^D\right) \\ &= \mathrm{cov}\left(\log \varepsilon_t^D + \log \varepsilon_t^W, \log \varepsilon_t^D\right) \\ &= \sigma_D^2 + \sigma_{DW}.\end{aligned}$$

The conditional correlation is

$$\begin{aligned}\rho_{c,d,t} &= corr_{t-1}\left(\log\frac{C_t}{C_{t-1}}, \log\frac{D_t}{D_{t-1}}\right)\\ &= \frac{\sigma_D^2+\sigma_{DW}}{\sqrt{\left(\sigma_D^2+\sigma_W^2+2\sigma_{DW}\right)}\sigma_D}\\ &= 7\frac{\sigma_D^2+\sigma_{DW}}{\sigma_D^2},\end{aligned}$$

where the last line uses (A9) and (A11). Targeting a correlation of 0.2 thus delivers

$$\sigma_{DW} = \left(\frac{0.2}{7}-1\right)\sigma_D^2. \tag{A12}$$

From (A11) we then get

$$\sigma_W^2 = \left(\frac{1}{49}-1-2\left(\frac{0.2}{7}-1\right)\right)\sigma_D^2. \tag{A13}$$

Equations (A12) and (A13) deliver our calibration targets. As is easily verified, the implied covariance matrix for the innovations $\left(\log\varepsilon_t^D, \log\varepsilon_t^W\right)$ has a positive determinant.

We now check what the calibration implies for the variance of unconditional log consumption growth σ_C^2. Using the fact that time t shocks are independent of $\log\left(1+\frac{W_{t-1}}{D_{t-1}}\right)$ and letting σ_C denote the standard deviation of unconditional log consumption growth, we get

$$\sigma_C^2 = \sigma_D^2 + 2\sigma_{DW} + \sigma_W^2 + (1-p)^2 \operatorname{var}\left(\log\left(1+\frac{W}{D}\right)\right).$$

The unconditional variance of $\log\left(1+\frac{W}{D}\right)$ is given by

$$\operatorname{var}\left(\log\left(1+\frac{W}{D}\right)\right) = \frac{\sigma_W^2}{1-p^2},$$

hence

$$\begin{aligned}\sigma_C^2 &= \sigma_D^2 + 2\sigma_{DW} + \frac{2(1-p)}{1-p^2}\sigma_W^2\\ &= \sigma_D^2 + 2\left(\frac{0.2}{7}-1\right)\sigma_D^2 + \frac{2(1-p)}{1-p^2}\left(\frac{1}{49}-1-2\left(\frac{0.2}{7}-1\right)\right)\sigma_D^2\end{aligned}$$

For $p \to 1$, we thus have the result that unconditional and conditional consumption growth volatility are identical ($\sigma_C = \sigma_{c,t} = \sigma_D/7$). For lower values of the persistence parameter p, unconditional consumption volatility decreases somewhat relative to conditional consumption volatility. For the parameter values considered in the main text, we have

$$\left(\frac{\sigma_D}{\sigma_C}\right) \approx \begin{cases} 7.0 & \text{for } p = 1.0 \\ 4.7 & \text{for } p = 0.95 \end{cases}.$$

A5. Existence of Optimum, Sufficiency of FOCs, Recursive Solution

Substituting out C_t using the budget constraint (14a) in problem (14), one obtains a problem that involves consumption choices only. Given the stock holding limits, the choice set is compact. It is also non-empty since $S_t = 1$ for all t is feasible. The following condition then insures existence of optimal plans:

Condition 1 The utility function $u(\cdot)$ is bounded above and for all $i \in [0, 1]$

$$E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t \, u\left(W_t + D_t\right) > -\infty. \tag{A14}$$

The expression on the left-hand side of condition (A14) is the utility associated with never trading stocks ($S_t^i = 1$ for all t). Since this policy is always feasible, condition (A14) guarantees that the objective function in (14a) is also bounded from below, even if the flow utility function $u(\cdot)$ is itself unbounded below.

From $\gamma > 1$, see equation (20), we have $u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \leq 0$ and thus a utility function that is bounded above. Provided (A14) holds, the optimization problem (14a) thus maximizes a bounded continuous utility function over a compact set, which guarantees existence of a maximum. Under the assumptions made in the main text (utility function given by (20), knowledge of (17) and (21)), condition (A14) indeed holds, as can be seen from the following derivations:

$$\begin{aligned} & E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t \, u\left(W_t + D_t\right) \\ = \; & E_0 \sum_{t=0}^{\infty} \delta^t \, u\left(W_t + D_t\right) \\ = \; & \frac{1}{1-\gamma} E_0 \sum_{t=0}^{\infty} \delta^t \left((1 + \frac{W_t}{D_t}) D_t \right)^{1-\gamma} \\ = \; & \frac{1}{1-\gamma} \left(\left(1 + \frac{W_0}{D_0}\right) D_0 \right)^{1-\gamma} E_0 \sum_{t=0}^{\infty} \delta^t \left(\frac{1 + \frac{W_t}{D_t}}{1 + \frac{W_0}{D_0}} \frac{D_t}{D_0} \right)^{1-\gamma}. \end{aligned}$$

Using

$$\begin{aligned} \frac{1 + \frac{W_t}{D_t}}{1 + \frac{W_0}{D_0}} &= \left(\frac{1+\rho}{1 + \frac{W_0}{D_0}} \right)^{1-p^t} \prod_{j=0}^{t-1} \left(\varepsilon_{t-j}^W\right)^{p^j} \\ \frac{D_t}{D_0} &= \left(\beta^D\right)^t \prod_{j=0}^{t-1} \varepsilon_{t-j}^D \end{aligned}$$

we get

$$
\begin{aligned}
& E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t \, u\left(W_t + D_t\right) \\
= \;& \frac{1}{1-\gamma}\left(\left(1+\frac{W_0}{D_0}\right) D_0\right)^{1-\gamma} E_0 \sum_{t=0}^{\infty}\left(\frac{1+\rho}{1+\frac{W_0}{D_0}}\right)^{\left(1-p^t\right)(1-\gamma)} \left(\delta \beta^D\right)^{t(1-\gamma)}\left(\prod_{j=0}^{t-1} \varepsilon_{t-j}^D \left(\varepsilon_{t-j}^W\right)^{p^j}\right)^{1-\gamma}
\end{aligned}
$$

The infinite sum is bounded if $\delta\left(\beta^D\right)^{1-\gamma} E[\left(\left(\varepsilon^W\right)^{p^j} \varepsilon^D\right)^{1-\gamma}]$ is bounded below one for all $j > 0$. The following derivations establish this fact:

$$
\begin{aligned}
\delta\left(\beta^D\right)^{1-\gamma} E\left[\left(\left(\varepsilon^W\right)^{p^j} \varepsilon^D\right)^{1-\gamma}\right] &= \delta\left(\beta^D\right)^{1-\gamma}\left(E\left[\left(\varepsilon^W\right)^{p^j(1-\gamma)}\right] + E[\left(\varepsilon^D\right)^{1-\gamma}]\right) \\
&= \delta\left(\beta^D\right)^{1-\gamma}\left(e^{p^j(\gamma-1)\left(1+p^j(\gamma-1)\right)\frac{\sigma_W^2}{2}} + e^{\gamma(\gamma-1)\frac{\sigma_D^2}{2}}\right) \\
&\leq \delta\left(\beta^D\right)^{1-\gamma}\left(e^{(\gamma-1)(1+(\gamma-1))\frac{\sigma_W^2}{2}} + e^{\gamma(\gamma-1)\frac{\sigma_D^2}{2}}\right) \\
&= \delta\left(\beta^D\right)^{1-\gamma} E\left[\left(\varepsilon^W\right)^{-\gamma}\left(\varepsilon^D\right)^{1-\gamma}\right] < 1,
\end{aligned}
$$

where the weak inequality follows from $\gamma > 1$ and $p \in [0,1]$ and the strict inequality from (21). This establishes existence of optimal plans.

Since (14a) is a strictly concave maximization problem the maximum is unique. With the utility function being differentiable, the first order conditions

$$
u'(C_t^i) = \delta E_t^{\mathcal{P}^i}\left[u'(C_{t+1}^i)\frac{P_{t+1}+D_{t+1}}{P_t}\right] \tag{A15}
$$

plus a standard transversality condition are necessary and sufficient for the optimum.

Recursive Formulation. We have a recursive solution whenever the optimal stock-holding policy can be written as a time-invariant function $S_t^i = S^i(x_t)$ of some state variables x_t. We seek a recursive solution where x_t contains appropriately rescaled variables that do not grow to infinity. With this in mind, we impose the following condition:

Condition 2 The flow utility function $u(\cdot)$ is homogeneous of degree $\eta \geq 0$. Furthermore, the beliefs $\mathcal{P}^i$ imply that $\theta_t \equiv \left(\frac{D_t}{D_{t-1}}, \frac{P_t}{D_t}, \frac{W_t}{D_t}\right)$ has a state space representation, i.e., the conditional distribution $\mathcal{P}^i(\theta_{t+1}|\omega^t)$ can be written as

$$
\begin{aligned}
\mathcal{P}^i(\theta_{t+1}|\omega^t) &= \mathcal{F}^i(m_t^i) && \text{(A16)} \\
m_t^i &= \mathcal{R}^i(m_{t-1}^i, \theta_t) && \text{(A17)}
\end{aligned}
$$

for some finite-dimensional state vector m_t^i and some time-invariant functions $\mathcal{F}^i$ and $\mathcal{R}^i$.

Under Condition 2, problem (14a) can then be re-expressed as

$$\text{(A18)} \qquad \max_{\{S_t^i \in [\underline{S}, \overline{S}]\}_{t=0}^{\infty}} E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t \, \mathcal{D}_t \; u\left(S_{t-1}^i \left(\frac{P_t}{D_t} + 1\right) - S_t^i \frac{P_t}{D_t} + \frac{W_t}{D_t}\right),$$

given $S_{-1}^i = 1$, where $\mathcal{D}_t$ is a time-varying discount factor satisfying $\mathcal{D}_{-1} = 1$ and

$$\mathcal{D}_t = \mathcal{D}_{t-1} \left(\beta^D \varepsilon_t^D\right)^{\eta} .$$

The return function in (A18) depends only on the exogenous variables contained in the vector θ_t. Since the beliefs $\mathcal{P}^i$ are assumed to be recursive in θ_t, standard arguments in dynamic programming guarantee that the optimal solution to (A18) takes the form (22). This formulation of the recursive solution is useful, because scaling P_t and W_t by the level of dividends eliminates the trend in these variables, as desired. This will be useful when computing numerical approximations to $S^i(\cdot)$. The belief systems $\mathcal{P}^i$ introduced in section V will satisfy the requirements stated in condition 2.

A6. *Proof of Proposition 2*

For general p we have

$$\begin{aligned} (1 + \frac{W_t}{D_t}) &= (1+\rho)^{(1-p)} (1 + \frac{W_{t-1}}{D_{t-1}})^p \ln \varepsilon_t^W \\ D_t &= \beta^D D_{t-1} \varepsilon_t^D \end{aligned}$$

so that for $S_t^i = 1$ for all $t \geq 0$, the budget constraint implies $C_t^i = D_t + W_t = D_t(1 + \frac{W_t}{D_t})$.

Substituting this into the agent's first order condition delivers

$$\text{(A19)} \qquad P_t = \delta E_t \left[\left(\frac{(1 + \frac{W_{t+1}}{D_{t+1}}) D_{t+1}}{(1 + \frac{W_t}{D_t}) D_t} \right)^{-\gamma} (P_{t+1} + D_{t+1}) \right].$$

Assuming that the following transversality condition holds

$$\text{(A20)} \qquad \lim_{j \to \infty} E_t \left[\delta^j \left\{ \left(\frac{1 + \frac{W_{t+j}}{D_{t+j}}}{1 + \frac{W_t}{D_t}} \right) \frac{D_{t+j}}{D_t} \right\}^{-\gamma} P_{t+j} \right] = 0,$$

one can iterate forward on (A19) to obtain

$$\frac{P_t}{D_t} = E_t \left[\sum_{j=1}^{\infty} \delta^j \left(\frac{1 + \frac{W_{t+j}}{D_{t+j}}}{1 + \frac{W_t}{D_t}} \right)^{-\gamma} \left(\frac{D_{t+j}}{D_t} \right)^{1-\gamma} \right].$$

Using $D_{t+j}/D_t = (\beta^D)^j \prod_{k=1}^{j} \varepsilon^D_{t+k}$ and

$$\frac{1+\frac{W_{t+j}}{D_{t+j}}}{1+\frac{W_t}{D_t}} = \left(\frac{1+\rho}{1+\frac{W_t}{D_t}}\right)^{1-p^j} \prod_{i=0}^{t-1} \left(\varepsilon^W_{t+j-i}\right)^{p^i},$$

one has

$$\frac{P_t}{D_t} = E_t\left[\sum_{j=1}^{\infty} \delta^j \left(\frac{1+\frac{W_{t+j}}{D_{t+j}}}{1+\frac{W_t}{D_t}}\right)^{-\gamma} \left(\frac{D_{t+j}}{D_t}\right)^{1-\gamma}\right]$$

$$\text{(A21)} \quad = \sum_{j=1}^{\infty} \left(\delta(\beta^D)^{1-\gamma}\right)^j E_t\left[\left(\frac{1+\rho}{1+\frac{W_t}{D_t}}\right)^{-\gamma(1-p^j)} \prod_{k=1}^{j} \left(\varepsilon^D_{t+k}\right)^{1-\gamma} \left(\varepsilon^W_{t+k}\right)^{-\gamma p^i}\right]$$

The infinite sum in the previous expression is bounded, if $E_t\left[\left(\varepsilon^W_{t+1}\right)^{-\gamma p^{j-1}} \left(\varepsilon^D_{t+1}\right)^{1-\gamma}\right]$ remains bounded away from one for all $j > 1$. This follows from the following derivations:

$$\begin{aligned}
& E_t\left[\left(\varepsilon^W_{t+1}\right)^{-\gamma p^{j-1}} \left(\varepsilon^D_{t+1}\right)^{1-\gamma}\right] = \\
= \; & E_t\left[\left(\varepsilon^W_{t+1}\right)^{-\gamma p^{j-1}}\right] + E_t\left[\left(\varepsilon^D_{t+1}\right)^{1-\gamma}\right] \\
= \; & e^{-\gamma p^{j-1}\frac{-\sigma^2_W}{2} + \left(\gamma p^{j-1}\right)^2 \frac{\sigma^2_W}{2}} + e^{(1-\gamma)\frac{-\sigma^2_D}{2} + (1-\gamma)^2\frac{\sigma^2_D}{2}} \\
\leq \; & e^{\gamma(1+\gamma)\frac{\sigma^2_W}{2}} + e^{(\gamma-1)\gamma\frac{\sigma^2_D}{2}} = E_t\left[\left(\varepsilon^W_{t+1}\right)^{-\gamma} \left(\varepsilon^D_{t+1}\right)^{1-\gamma}\right] < 1,
\end{aligned}$$

where the inequality in the third to last line follows from $\gamma > 1$ and $p \in [0, 1]$ and the last inequality from assumption (21). For the special cases $p = 1$ and $p = 0$, equation (A21) delivers the expressions stated in proposition 2.

A7. Bayesian Foundations for Lagged Belief Updating

We now present a slightly modified information structure for which Bayesian updating gives rise to the lagged belief updating equation (39). Specifically, we generalize the perceived price process (26) by splitting the temporary return innovation $\ln \varepsilon_{t+1}$ into two independent subcomponents:

$$\ln P_{t+1} - \ln P_t = \ln \beta_{t+1} + \ln \varepsilon^1_{t+2} + \ln \varepsilon^2_{t+1}$$

with $\ln \varepsilon^1_{t+2} \sim iiN(-\frac{\sigma^2_{\varepsilon,1}}{2}, \sigma^2_{\varepsilon 1})$, $\ln \varepsilon^2_{t+1} \sim iiN(-\frac{\sigma^2_{\varepsilon 2}}{2}, \sigma^2_{\varepsilon 2})$ and

$$\sigma^2_\varepsilon = \sigma^2_{\varepsilon 1} + \sigma^2_{\varepsilon 2}.$$

We then assume that in any period t agents observe the prices, dividends and wages up to period t, as well as the innovations ε_t^1 up to period t. Agents' time t information set thus consists of $I_t = \{P_t, D_t, W_t, \varepsilon_t^1, P_{t-1}, D_{t-1}, W_{t-1}, \varepsilon_{t-1}^1, ...\}$. By observing the innovations ε_t^1, agents learn - with a one period lag - something about the temporary components of price growth. The process for the persistent price growth component $\ln \beta_t$ remains as stated in equation (27), but we now denote the innovation variance by $\sigma_{\widetilde{v}}^2$ instead of σ_v^2. As before, $\ln m_t$ denotes the posterior mean of $\ln \beta_t$ given the information available at time t. We prove below the following result:

PROPOSITION 4: *Fix $\sigma_\varepsilon^2 > 0$ and consider the limit $\sigma_{\varepsilon 2}^2 \to 0$ with $\sigma_{\widetilde{v}}^2 = \sigma_{\varepsilon 2}^2 g^2/(1-g)$. Bayesian updating then implies*

$$\ln m_t = \ln m_{t-1} + g\left(\ln P_{t-1} - \ln P_{t-2} - \ln m_{t-1}\right) - g \ln \varepsilon_t^1 \tag{A22}$$

The modified information structure thus implies that only lagged price growth rates enter the current state estimate, so that beliefs are predetermined, precisely as assumed in equation (39). Intuitively, this is so because lagged returns become infinitely more informative relative to current returns as $\sigma_{\varepsilon 2}^2 \to 0$, which eliminates the simultaneity problem. For non-vanishing uncertainty $\sigma_{\varepsilon 2}^2$ the weight of the last observation actually remains positive but would still be lower than that given to the lagged return observation, see equation (A25) in the proof below and the subsequent discussion for details.

We now provide the proof of the previous proposition. Let us define the following augmented information set $\widetilde{I}_{t-1} = I_{t-1} \cup \{\varepsilon_t^1\}$. The posterior mean for β_t given $\widetilde{I}_{t-1}$, denoted $\ln m_{t|\widetilde{I}_{t-1}}$ is readily recursively determined via

(A23)

$$\ln m_{t|\widetilde{I}_{t-1}} = \ln m_{t-1|\widetilde{I}_{t-2}} - \frac{\sigma_{\widetilde{v}}^2}{2} + \widetilde{g}\left(\ln P_{t-1} - \ln P_{t-2} - \ln \varepsilon_t^1 + \frac{\sigma_{\widetilde{v}}^2 + \sigma_{\varepsilon 2}^2}{2} - \ln m_{t-1|\widetilde{I}_{t-2}}\right)$$

and the steady state posterior uncertainty and the Kalman gain by

$$\begin{aligned} \sigma^2 &= \frac{-\sigma_{\widetilde{v}}^2 + \sqrt{\left(\sigma_{\widetilde{v}}^2\right)^2 + 4\sigma_{\widetilde{v}}^2 \sigma_{\varepsilon 2}^2}}{2} \\ \widetilde{g} &= \frac{\sigma^2}{\sigma_{\varepsilon 2}^2} \end{aligned} \tag{A24}$$

We furthermore have

$$E[\ln P_t - \ln P_{t-1} | \widetilde{I}_{t-1}] = \ln m_{t|\widetilde{I}_{t-1}} - \frac{\sigma_{\varepsilon 1}^2 + \sigma_{\varepsilon 2}^2}{2}$$

and

$$(\ln P_t - \ln P_{t-1}) - E[\ln P_t - \ln P_{t-1} | \widetilde{I}_{t-1}] = \ln \beta_t - \ln m_{t|\widetilde{I}_{t-1}} + \ln \varepsilon_{t+1}^1 + \ln \varepsilon_t^2 + \frac{\sigma_{\varepsilon 1}^2 + \sigma_{\varepsilon 2}^2}{2}$$

so that

$$\left(\begin{array}{c} \ln \beta_t \\ \ln P_t - \ln P_{t-1} \end{array} |\widetilde{I}_{t-1} \right) \sim N\left(\left(\begin{array}{c} \ln m_{t|\widetilde{I}_{t-1}} \\ \ln m_{t|\widetilde{I}_{t-1}} - \frac{\sigma^2_{\varepsilon 1}+\sigma^2_{\varepsilon 2}}{2} \end{array} \right), \left(\begin{array}{cc} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + \sigma^2_{\varepsilon 1} + \sigma^2_{\varepsilon 2} \end{array} \right) \right),$$

where the covariance between $\ln \beta_t | \widetilde{I}_{t-1}$ and $\ln P_t - \ln P_{t-1} | \widetilde{I}_{t-1}$ can be computed by exploiting the fact that $\ln \beta_t - \ln m_{t|\widetilde{I}_{t-1}}$ and $\ln \varepsilon^1_{t+1} + \ln \varepsilon^2_t$ are independent and using $\ln \beta_t - E_t[\ln \beta_t | \widetilde{I}_{t-1}] = \ln \beta_t - \ln m_{t|\widetilde{I}_{t-1}}$. Using standard normal updating formulae, we can thus compute

$$\begin{aligned} \ln m_{t|I_t} &= E[\ln \beta_t | I_t] = E[\ln \beta_t | \widetilde{I}_{t-1}, \ln P_t - \ln P_{t-1}] \\ \text{(A25)} = \quad & \ln m_{t|\widetilde{I}_{t-1}} + \frac{\sigma^2}{\sigma^2 + \sigma^2_{\varepsilon 1} + \sigma^2_{\varepsilon 2}} \left(\ln P_t - \ln P_{t-1} + \frac{\sigma^2_{\varepsilon 1} + \sigma^2_{\varepsilon 2}}{2} - \ln m_{t|\widetilde{I}_{t-1}} \right), \end{aligned}$$

where the second equality exploits the fact that D_t, W_t contain no information about $\ln \beta_t$, and the second inequality follows from the fact that $\widetilde{I}_{t-1}$ contains information up to $t-1$ and including ε^1_t, since the latter is independent of $(\ln P_t - \ln P_{t-1})$.

Since $\frac{\sigma^2}{\sigma^2+\sigma^2_{\varepsilon 1}+\sigma^2_{\varepsilon 2}} < \frac{\sigma^2}{\sigma^2_{\varepsilon 2}} = \widetilde{g}$, the weight of the price observation dated t is reduced relative to the earlier observation dated $t-1$ because it is ‘noisier’. Now consider the limit $\sigma^2_{\varepsilon 2} \to 0$ and along the limit choose $\sigma^2_{\varepsilon 1} = \sigma^2_{\varepsilon} - \sigma^2_{\varepsilon 2}$ and $\sigma^2_{\widetilde{\upsilon}} = \frac{g^2}{1-g}\sigma^2_{\varepsilon 2}$, as assumed in the proposition. From $\sigma^2 \to 0$ and equation (A25) it then follows that $\ln m_{t|I_t} = \ln m_{t|\widetilde{I}_{t-1}}$, i.e., the weight of the last observation price converges to zero. Moreover, from $\sigma^2_{\widetilde{\upsilon}} = \frac{g^2}{1-g}\sigma^2_{\varepsilon 2}$ and (A24) we get $\widetilde{g} = g$. Using these results, equation (A23) implies equation (A22).

A8. Proof of Proposition 3

The proof uses the assumption of no uncertainty so that for any function f we have $E^{\mathcal{P}}_t f(X_{t+j}, Y_{t+j}) = f(E^{\mathcal{P}}_t X_{t+j}, E^{\mathcal{P}}_t Y_{t+j})$. Simplifying notation (and slightly abusing it) in this appendix we let $X_{t+j} = E^{\mathcal{P}}_t X_{t+j}$ for all $j \geq 1$, so that X_{t+j} below denotes the subjective expectation conditional on information at time t of the variable X at time $t+j$. The first order conditions (A15) can then be written as

$$\text{(A26)} \qquad C_t^{-\gamma} = C_{t+1}^{-\gamma} \delta R_{t+1} \Longrightarrow C_t = \delta^{-\frac{j}{\gamma}} \prod_{\tau=1}^{j} R_{t+\tau}^{-\frac{1}{\gamma}} C_{t+j}$$

for all t, $j \geq 0$, assuming the stock limits are not binding in periods $t, t+1, ..., t+j-1$. Iterating forward N periods on the budget constraint of the agent and using the fact that either $\prod_{\tau=1}^{N} R_{t+\tau}^{-1} \to 0$ or $S_{t+N} \to 0$ as $N \to \infty$ we have

$$(P_t + D_t)\, S_{t-1} = \sum_{j=0}^{\infty} \left(\prod_{\tau=1}^{j} R_{t+\tau}^{-1} \right) \left(C_{t+j} - W_{t+j} \right)$$

Using equation (A26) to substitute out C_{t+j} gives

$$(P_t + D_t)\, S_{t-1} = \sum_{j=0}^{\infty} \left(\prod_{\tau=1}^{j} R_{t+\tau}^{-1} \right) \left[\left(\frac{W_t}{D_t} + 1 \right) \delta^{\frac{j}{\gamma}} \prod_{\tau=1}^{j} R_{t+\tau}^{\frac{1}{\gamma}} - W_{t+j} \right]$$

assuming the stock limits are large enough not to be binding.

Imposing on the previous equation $S_{t-1} = 1$ (the market clearing condition for S_{t-1} if $t \geq 0$, or the initial condition for period $t = 0$) and $C_t = D_t + W_t$ (the market clearing condition for consumption) one obtains

$$\frac{P_t}{D_t} + 1 = \sum_{j=0}^{N} \left(\prod_{\tau=1}^{j} R_{t+\tau}^{-1} \right) \left[\left(\frac{W_t}{D_t} + 1 \right) \delta^{\frac{j}{\gamma}} \prod_{\tau=1}^{j} R_{t+\tau}^{\frac{1}{\gamma}} - \frac{W_{t+j}}{D_t} \right]$$

Cancelling the terms for $j = 0$ in each summation gives for the market-clearing price

$$\frac{P_t}{D_t} = \sum_{j=1}^{N} \left(\prod_{\tau=1}^{j} R_{t+\tau}^{-1} \right) \left[\left(\frac{W_t}{D_t} + 1 \right) \delta^{\frac{j}{\gamma}} \prod_{\tau=1}^{j} R_{t+\tau}^{\frac{1}{\gamma}} - \frac{W_{t+j}}{D_t} \right]$$

Using (40) gives (41).

A9. Verification of Conditions (40)

For the vanishing noise limit of the beliefs specified in section V we have

$$\begin{aligned} E_t^{\mathcal{P}}[P_{t+j}] &= (m_t)^j\, P_t \\ E_t^{\mathcal{P}}[D_{t+j}] &= \left(\beta^D\right)^j D_t \\ E_t^{\mathcal{P}}[W_{t+j}] &= \left(\beta^D\right)^j W_t \end{aligned}$$

where we have abstracted from transitional dynamics in the W_t/D_t ratio and assume $W_t/D_t = \rho$, as transitional dynamics do not affect the limit results. We first verify the inequality on the l.h.s. of equation (40). We have

$$\lim_{T\to\infty} E_t^{\mathcal{P}}[R_T] = m_t + \lim_{T\to\infty} \left(\frac{\beta^D}{m_t} \right)^{T-1} \beta^D \frac{D_t}{P_t},$$

so that for $m_t > 1$ the limit clearly satisfies $\lim_{T\to\infty} E_t^{\mathcal{P}}[R_T] > 1$ due to the first term on the r.h.s.; for $m_t < 1$ the second term on the r.h.s. increases without bound, due to $\beta^D > 1$, so that $\lim_{T\to\infty} E_t^{\mathcal{P}}[R_T] > 1$ also holds.

In a second step we verify that the inequality condition on the r.h.s. of equation (40) holds for all subjective beliefs $m_t > 0$. We have

$$\lim_{T\to\infty} E_t^{\mathcal{P}}\left(\sum_{j=1}^{T}\left(\prod_{i=1}^{j}\frac{1}{R_{t+i}}\right)W_{t+j}\right) = \lim_{T\to\infty} W_t E_t^{\mathcal{P}}\left(\sum_{j=1}^{T}\left(\beta^D\right)^j\left(\prod_{i=1}^{j}\frac{1}{R_{t+i}}\right)\right)$$

$$= \lim_{T\to\infty} W_t \sum_{j=1}^{T} X_j \tag{A27}$$

where

$$X_j = \frac{\left(\beta^D\right)^j}{\prod_{i=1}^{j}(m_t + \left(\frac{\beta^D}{m_t}\right)^{i-1}\beta^D\frac{D_t}{P_t})} \geq 0 \tag{A28}$$

A sufficient condition for the infinite sum in (A27) to converge is that the terms X_j are bounded by some exponentially decaying function. The denominator in (A28) satisfies

$$\prod_{i=1}^{j}(m_t + \left(\frac{\beta^D}{m_t}\right)^{i-1}\beta^D\frac{D_t}{P_t})$$

$$\geq (m_t)^j + \left(\frac{\beta^D}{m_t}\right)^{j\left(\frac{j-1}{2}\right)}\beta^D\frac{D_t}{P_t}, \tag{A29}$$

where the first term captures the the pure products in m_t, the second term the pure products in $\left(\frac{\beta^D}{m_t}\right)^{i-1}\beta^D\frac{D_t}{P_t}$, and all cross terms have been dropped. We then have

$$\begin{aligned} X_j &= \frac{\left(\beta^D\right)^j}{\prod_{i=1}^{j}(m_t + \left(\frac{\beta^D}{m_t}\right)^{i-1}\beta^D\frac{D_t}{P_t})} \\ &\leq \frac{\left(\beta^D\right)^j}{(m_t)^j + \left(\frac{\beta^D}{m_t}\right)^{j\left(\frac{j-1}{2}\right)}\beta^D\frac{D_t}{P_t}} \\ &= \frac{1}{\left(\frac{m_t}{\beta^D}\right)^j + \left(\frac{\beta^D}{m_t}\right)^{j\left(\frac{j-1}{2}\right)}\frac{1}{\left(\beta^D\right)^{j-1}}\frac{D_t}{P_t}}, \end{aligned}$$

where all terms in the denominator are positive. For $m_t \geq \beta^D > 1$ we can use the first term in the denominator to exponentially bound X_j, as $X_j \leq \left(\frac{\beta^D}{m_t}\right)^j$; for $m_t < \beta^D$ we can use the second term:

$$X_j \leq \frac{1}{\left(\frac{\beta^D}{m_t}\right)^{j\left(\frac{j-1}{2}\right)} \frac{1}{(\beta^D)^{j-1}} \frac{D_t}{P_t}} = \frac{1}{\left(\left(\frac{\beta^D}{m_t}\right)^{\frac{j}{2}} \frac{1}{\beta^D}\right)^{j-1} \frac{D_t}{P_t}}$$

Since $m_t < \beta^D$ there must be a $J < \infty$ such that

$$\left(\frac{\beta^D}{m_t}\right)^{\frac{j}{2}} \frac{1}{\beta^D} \geq \frac{\beta^D}{m_t} > 1$$

for all $j \geq J$, so that the X_j are exponentially bounded for all $j \geq J$.

A10. Proof of Lemma 1

Proof of lemma 1: We start by proving the first point in the lemma. The price, dividend and belief dynamics in the deterministic model are described by the following equations

$$\begin{aligned} \ln m_t &= \ln m_{t-1} + g\left(\ln P_{t-1} - \ln P_{t-2} - \ln m_{t-1}\right) \\ \ln P_t - \ln D_t &= f(\ln m_t) \\ \ln D_t - \ln D_{t-1} &= \beta^D, \end{aligned} \tag{A30}$$

where $f(\cdot)$ is a continuos function, implicitly defined by the log of the P_t/D_t solution to equation (42).[100] Substituting the latter two equations into the first delivers

$$\ln m_t - \ln m_{t-1} = g\left(f(\ln m_{t-1}) - f(\ln m_{t-2}) + \ln \beta^D - \ln m_{t-1}\right).$$

If $\ln m_t$ converges, then the l.h.s. of the previous equation must converge to zero. Since $f(\cdot)$ is continuos, this means that m_{t-1} must converge to β^D, as claimed.

We now prove the second point in the lemma. The belief dynamics implied by the second order difference equation (A30) can expressed as a two-dimensional first order difference equation using the mapping $F : R^2 \rightarrow R^2$, defined as

$$F(x) = \begin{pmatrix} x_1 + g(\ln PD(e^{x_1}) - \ln PD(e^{x_2}) + \ln \beta^D - x_1) \\ x_1 \end{pmatrix},$$

so that

$$\begin{pmatrix} \ln m_t \\ \ln m_{t-1} \end{pmatrix} = F\begin{pmatrix} \ln m_{t-1} \\ \ln m_{t-2} \end{pmatrix}.$$

Clearly, F has a fixed point at the RE solution, i.e., $(\ln \beta^D, \ln \beta^D)' = F(\ln \beta^D, \ln \beta^D)'$. Moreover, m_t locally converges to the RE beliefs if and only if

$$\frac{\partial F(\ln \beta^D, \ln \beta^D)}{\partial x'} = \begin{pmatrix} 1 + g\,(\xi - 1) & -g\,\xi \\ 1 & 0 \end{pmatrix} \tag{A31}$$

[100] Since we are interested in asymptotic results and since $W_t/D_t \rightarrow \rho$, pricing is asymptotically given by equation (42).

has all eigenvalues less than one in absolute value, where $\xi \equiv \left.\frac{\partial \ln PD(e^{\ln m})}{\partial \ln m}\right|_{m=\beta^D} = \frac{\beta^D}{PD(\beta^D)} \left.\frac{\partial PD(m)}{\partial m}\right|_{m=\beta^D}$ with $|\xi| < 1$. The eigenvalues of the matrix in equation (A31) are

$$\lambda = \frac{1 + g(\xi - 1) \pm \sqrt{(1 + g(\xi - 1))^2 - 4g\,\xi}}{2}.$$

From $|\xi| < 1$ and $g < \frac{1}{2}$ follows that $(1 + g(\xi - 1))^2 - 4g\xi > 1 - 2g\xi - 2g \geq 0$, so that all eigenvalues are real. As is easily verified, we have $\lambda^+ < 1$ because

$$\begin{array}{rcl} 1 + g(\xi - 1) & < & 2 - \sqrt{(1 + g(\xi - 1))^2 - 4g\,\xi} \Leftrightarrow \\ \sqrt{(1 + g(\xi - 1))^2 - 4g\,\xi} & < & 1 - g(\xi - 1) \Leftrightarrow \\ (1 + g(\xi - 1))^2 - 4g\,\xi & < & (1 - g(\xi - 1))^2 \Leftrightarrow \\ 2g(\xi - 1) - 4g\xi & < & -2g(\xi - 1) \Leftrightarrow \\ -g & < & 0 \end{array}$$

and $\lambda^- < 1$ because

$$-1 + g(\xi - 1) < \sqrt{(1 + g(\xi - 1))^2 - 4g\,\xi}$$

where the l.h.s. is negative and the r.h.s. positive. We have $\lambda^+ > -1$ if and only if

$$1 + g(\xi - 1) > -2 - \sqrt{(1 + g(\xi - 1))^2 - 4g\,\xi}$$

From $|\xi| < 1$ and $g < \frac{1}{2}$ the l.h.s. is weakly positive, while the r.h.s. is strictly negative. We have $\lambda^- > -1$ if and only if

$$\begin{array}{rcl} 3 + g(\xi - 1) & > & \sqrt{(1 + g(\xi - 1))^2 - 4g\,\xi} \Leftrightarrow \\ (2 + (1 + g(\xi - 1)))^2 & > & (1 + g(\xi - 1))^2 - 4g\,\xi \Leftrightarrow \\ 1 + (1 + g(\xi - 1)) & > & -g\,\xi \Leftrightarrow \\ 2 + g(2\xi - 1) & > & 0 \end{array}$$

The last equation holds since $|\xi| < 1$ and $g < \frac{1}{2}$. This shows that the eigenvalues of (A31) are all inside the unit circle.

A11. Capital Gains Expectations and Expected Returns: Further Details

Figure A1 depicts how expected returns at various horizons depend on agent's expected price growth expectations using the same parameterization as used in figure 5. It shows that expected returns covary positively with capital gains expectations for $m_t \geq \beta^D$, as has been claimed in the main text. The flat part at around $m_t - 1 \approx 0.01$ arises because in that area the PD ratio increases strongly, so that the dividend yield falls. Only

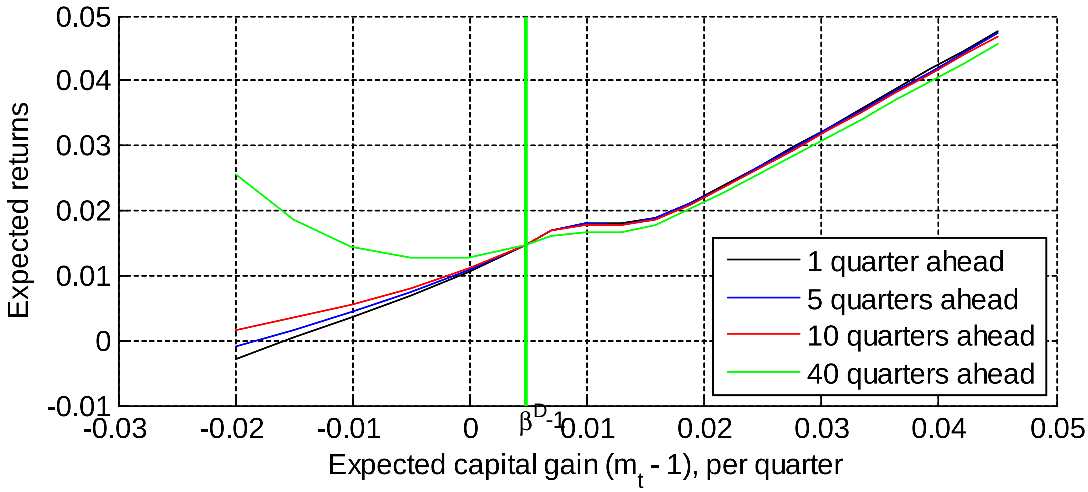

FIGURE A1. EXPECTED RETURN AS A FUNCTION OF EXPECTED CAPITAL GAIN

for pessimistic price growth expectations ($m_t < \beta^D$) and long horizons of expected returns we find a negative relationship. The latter emerges because with prices expected to fall, the dividend yield will rise and eventually result in high return expectations.

A12. Numerical Solution Algorithm

Algorithm: We solve for agents' state-contingent, time-invariant stockholdings (and consumption) policy (22) using time iteration in combination with the method of endogenous grid points. Time iteration is a computationally efficient, e.g., Aruoba et al. 2006, and convergent solution algorithm, see Rendahl 2013. The method of endogenous grid points, see Carroll 2006, economizes on a costly root finding step which speeds up computations further.

Evaluations of Expectations: Importantly, agents evaluate the expectations in the first order condition (A15) according to their subjective beliefs about future price growth and their (objective) beliefs about the exogenous dividend and wage processes. Expectations are approximated via Hermite Gaussian quadrature using three interpolation nodes for the exogenous innovations.

Approximation of Optimal Policy Functions: The consumption/stockholding policy is approximated by piecewise linear splines, which preserves the nonlinearities arising in particular in the PD dimension of the state space. Once the state-contingent consumption policy has been found, we use the market clearing condition for consumption goods to determine the market clearing PD ratio for each price-growth belief m_t.

Accuracy: Carefully choosing appropriate grids for each belief is crucial for the accuracy of the numerical solution. We achieve maximum (relative) Euler errors on the order of 10^{-3} and median Euler errors on the order of 10^{-5} (average: 10^{-4}).

Using our analytical solution for the case with vanishing noise, we can assess the accuracy of our solution algorithm more directly. Setting the standard deviations of exogenous disturbances to 10^{-16} the algorithm almost perfectly recovers the equilibrium PD ratio of the analytical solution: the error for the numerically computed equilibrium PD ratio for any price growth belief m_t on our grid is within 0.5 % of the analytical solution.

A13. Model with Stock Supply Shocks

With the supply shocks specified in the main text, the individual optimization problem remains unchanged. The only point where the model changes is when we compute market clearing prices. These now have to satisfy the relation

$$S(e^{\varepsilon^s_{t-1}}, \frac{P_t}{D_t}, \frac{W_t}{D_t}, m_t) = e^{\varepsilon^s_t},$$

where

$$\varepsilon^s_t \sim iiN(-\frac{\sigma^2_{\varepsilon^s}}{2}, \sigma^2_{\varepsilon^S}).$$

To illustrate model performance in the presence of supply shocks and to compare it to the model without supply shocks, we continue to keep the parameter values for (γ, δ, p) equal to the estimated ones for the model from table 3 (diagonal matrix) and only reestimate the gain value g, so as to fit the asset pricing moments reported in table 3, using a diagonal weighting matrix. Clearly, an even better match with the data moments can be achieved by reestimating all parameters. For $\sigma^2_{\varepsilon^S} = 1.25 \cdot 10^{-3}$, the new gain estimate is $\widehat{g} = 0.02315$.

Table A5 below reports the moments and t-ratios for the model with supply shocks. The model implied standard deviation for the risk free rate and the autocorrelations of the excess stock returns and stock returns are reported in table 6 in the main text. In terms of the t-ratios in table A1, the model performs equally well as the baseline models from table 3, except for the autocorrelation of the PD ratio which is now somewhat lower. The fit with the latter moment could be improved further by relaxing the assumption that ε^s_t

is iid.

	Subj. Belief Model with Supply Shocks	
	Moment	t-ratio
E[PD]	107.3	-1.30
Std[PD]	74.1	0.60
Corr[$\mathbf{PD}_t$, $\mathbf{PD_{t-1}}$]	0.96	-6.66
Std[R]	7.45	-1.36
c	-0.0044	-0.28
R^2	0.16	-0.63
E[R]-1	1.81	-0.18
E[R^b]-1	0.98	4.99
UBS Survey Data:		
Corr[$\mathbf{PD}_t$,$\mathbf{E}_t^{\mathcal{P}}\mathbf{R}_{t,t+4}$]	0.75	-0.69

Table A5: Asset pricing moments, estimated model with supply shocks from table 6 in the main text

A14. Derivation of Approximate Sharpe Ratios

Under rational and subjective price beliefs, the following two first-order-conditions hold, namely the first order condition for stocks

$$1 = E_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}(1+r_{t+1}^s)\right], \tag{A32}$$

where $1+r_{t+1}^s = \frac{P_{t+1}+D_{t+1}}{P_t}$ is the gross real stock return, and the first order condition for bonds

$$1+r_t^b = \frac{1}{E_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]}. \tag{A33}$$

Equation (A32) can be written as

$$E_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]\cdot E_t^{\mathcal{P}}\left[r_{t+1}^s\right] + cov_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}, r_{t+1}^s\right] = 1. \tag{A34}$$

Dividing the previous equation by $E_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]$ and using (A33) one obtains

$$E_t^{\mathcal{P}}\left[r_{t+1}^s\right]-r_t^b=-(1+r_t^b)\cdot cov_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma},1+r_{t+1}^s\right]. \tag{A35}$$

Dividing by $Std_t^{\mathcal{P}}\left[r_{t+1}^s\right]$ delivers

$$\frac{E_t^{\mathcal{P}}\left[r_{t+1}^s\right]-r_t^b}{Std_t^{\mathcal{P}}\left[r_{t+1}^s\right]}=-(1+r_t^b)\cdot Std_t^{\mathcal{P}}\left[\delta\left(C_{t+1}/C_t\right)^{-\gamma}\right]corr_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma},1+r_{t+1}^s\right], \tag{A36}$$

where the left-hand side is the (subjective) conditional Sharpe ratio and $corr_t^{\mathcal{P}}[\cdot]$ the (subjective) conditional correlation. Using the fact that

$$\begin{aligned}\left(1+r_t^b\right)\delta &\approx 1\\ corr_t^{\mathcal{P}}\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma},1+r_{t+1}^s\right] &\approx -1,\end{aligned}$$

where the latter follows from the first order condition (A32), we have under the additional assumption of log-normal consumption growth (which holds exactly in the case with rational price expectations):

$$\frac{E_t^{\mathcal{P}}\left[r_{t+1}^s\right]-r_t^b}{Std_t^{\mathcal{P}}\left[r_{t+1}^s\right]}\approx\gamma\, Std_t^{\mathcal{P}}\left[C_{t+1}/C_t\right]. \tag{A37}$$

For the case with rational price expectations ($E_t^{\mathcal{P}}[\cdot]=E_t[\cdot]$, $Std_t^{\mathcal{P}}[\cdot]=Std_t[\cdot]$), it then follows from (24) that $Std_t\left[r_{t+1}^s\right]\approx Std\left[r_{t+1}^s\right]$, so that by using this relationship to substitute $Std_t\left[r_{t+1}^s\right]$ in (A37) and by applying the unconditional rational expectations operator on both sides of the equation, one obtains equation (45) in the main text.

For the case with subjective price expectations, we have

$$\begin{aligned}Std_t^{\mathcal{P}}\left[r_{t+1}^s\right] &\approx Std^{\mathcal{P}}\left[r_{t+1}^s\right]\\ &\approx Std\left[r^s\right],\end{aligned} \tag{A38}$$

where the first approximation is a feature of the subjective price belief system[101] and

[101] According to agents' subjective beliefs

$$\begin{aligned}1+r_{t+1}^s &= \frac{P_{t+1}+D_{t+1}}{P_t}\\ &= \beta_{t+1}\upsilon_{t+1}\varepsilon_{t+1}+\frac{D_t}{P_t}\varepsilon_t^D.\end{aligned} \tag{A39}$$

Since $\frac{D_t}{P_t}$ is small, $\beta_{t+1}\upsilon_{t+1}\approx 1$, and the standard deviation of ε_t^D is small relative to the standard deviation of ε_{t+1}, the last term in (A39) contributes little to the standard deviation of r_{t+1}^s. It then follows from $\beta_{t+1}\upsilon_{t+1}\approx 1$ that the second approximation due to the way we calibrated the standard deviation σ_ε of the transitory price shock ε_t in table 2. Using (A38) to substitute $Std_t^{\mathcal{P}}\left[r_{t+1}^s\right]$ in (A37) and applying the unconditional expectations operator on both sides of the equation delivers

$$\frac{E\left[E_t^{\mathcal{P}}\left[r_{t+1}^s\right]\right]-E\left[r_t^b\right]}{Std\left[r^s\right]}\approx\gamma\, E\left[Std_t^{\mathcal{P}}\left[C_{t+1}/C_t\right]\right],$$

which implies (46).

A15. Simultaneous Belief Updating

This appendix provides further information about the extended model with a share of simultaneous belief updaters considered in section X.C.

Table A6 reports for different values of α the model moments when the selection rule chooses the closest market clearing price.[102] The table shows that most asset pricing implications of the subjective belief model are rather robust to allowing for contemporaneous belief updating. The main quantitative effects of introducing current updaters consists of increasing the volatility of stock returns, the volatility of the PD ratio and the equity premium. These effects become more pronounced as the share of current updaters α increases, as this leads to an increase in the percentage share of periods with multiple market clearing prices. Importantly, however, the objectively expected discounted utility of agents that use lagged belief updating exceeds - for all values of α - that of agents who use current belief updating, see the last two rows in table A6.[103] Current updaters would thus have an incentive to switch to lagged updating, i.e., to the setting considered in our baseline specification.

Table A7 reports the model moments when the equilibrium selection rule chooses instead the market clearing price that is furthest away from the previous period's price. For $\alpha\leq 0.7$ the same phenomena occur as for the alternative selection rule considered before, i.e., volatility and risk premia increase, with the quantitative effects being now more pronounced. For $\alpha\geq 0.8$, the share of periods with multiple equilibria increases substantially, so that there are more often two consecutive periods with multiple equilibria. The selection rule then creates an oscillating pattern between high and low market clearing prices. This manifests itself in a reduced autocorrelation for the PD ratio, which becomes even negative for $\alpha\geq 0.9$. As before, current forecasters' (objectively) expected utility always falls short of that experienced by forecasters using only lagged price growth. For large values of α, the utility gap widens significantly because the forecast quality of forecasters using current price information deteriorates significantly in the presence of oscillating price patterns.

Overall, we find that - in line with the postulated subjective belief structure in our baseline setting - agents will find it optimal to use only lagged price observations to update beliefs. Even if some agents would use current price information for updating beliefs,

$Std_t^{\mathcal{P}}\left[r_{t+1}^s\right]\approx Std^{\mathcal{P}}\left[\varepsilon_{t+1}\right]$, which is time invariant, as claimed.

[102] The parameters in tables A6 and A7 are those given by the estimated model from table 3 (diagonal matrix). To compute $Corr[PD_t, E_t^{\mathcal{P}}[R_{t+1}]]$ we let $E_t^{\mathcal{P}}[R_{t+1}]$ denote the average return expectations across agents, in line with how the return expectations are computed in the survey data.

[103] The table reports the unconditional expectation of discounted consumption utility using the objective distribution for consumption, as realized in equilibrium.

the model continues to produce high amounts of stock price volatility and also tends to deliver a positive correlation between the PD ratio and subjective expected returns.

	U.S. Data	**Share of current updaters (α)**										
	1946:1-2012:1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
E[PD]	139.8	115.2	115.7	116.2	117.4	118.9	119.6	123.2	127.1	132.4	135.2	137.1
Std[PD]	65.2	88.2	89.3	90.5	92.9	95.9	97.2	103.2	108.7	115.8	118.8	121.7
Corr[PD_t, PD_{t-1}]	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.96	0.96	0.96	0.94
Std[R]	8.00	7.74	8.11	8.59	9.25	9.92	10.46	11.38	12.64	15.41	20.79	28.71
c	-0.0041	-0.0050	-0.0050	-0.0051	-0.0052	-0.0052	-0.0053	-0.0055	-0.0059	-0.0062	-0.0066	-0.0067
R^2	0.25	0.20	0.19	0.19	0.19	0.19	0.19	0.20	0.21	0.22	0.23	0.24
E[R]-1	1.89	1.82	1.85	1.89	1.94	2.01	2.08	2.22	2.41	2.73	3.22	3.96
E[$R^b - 1$]	0.13	0.99	0.96	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.94	1.03
UBS Survey Data:												
Corr[PD_t,$E_t^{\mathcal{P}} R_{t,t+4}$]	0.79	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.85	0.86	0.87	0.87
Percent of periods w												
multiple equilibria:		0.00	0.00	0.41	1.88	3.74	5.50	8.22	24.22	66.38	87.35	97.76
Expected utility:												
Lagged updating		-4.02	-4.02	-4.01	-4.01	-4.01	-4.00	-4.00	-3.99	-3.99	-3.97	-3.95
Current updating		-4.07	-4.06	-4.06	-4.05	-4.05	-4.05	-4.04	-4.04	-4.03	-4.03	-4.02

Table A6: Asset pricing moments with simultaneous belief updating
(market clearing price closest to previous period price)

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, see table 3 for a description of the labels used in the first column, and the asset pricing moments for the estimated model from table 3 (diagonal matrix), considering different values for the share α of belief updaters using current price growth observations (columns 3 to 13). The table also reports the percentage share of periods in which multiple equilibria are encountered along the equilibrium (fifth to last line) and the unconditional expected utility of agents using current or lagged price growth observations for updating beliefs (last two lines). In periods with multiple equilibria, the equilibrium price closest to the previous market clearing price is selected.

	U.S. Data	Share of current updaters (α)										
	1946:1-2012:1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
E[PD]	139.8	115.2	115.7	115.3	112.7	109.1	106.3	107.8	117.7	136.0	167.5	225.3
Std[PD]	65.2	88.2	89.3	89.2	84.9	78.0	71.5	75.5	92.2	111.9	124.2	177.3
Corr[PD$_t$, PD$_{t-1}$]	0.98	0.98	0.98	0.98	0.97	0.96	0.94	0.91	0.78	0.31	-0.51	-0.91
Std[R]	8.00	7.74	8.11	8.66	9.37	10.16	11.23	14.72	27.80	64.04	186.20	362.66
c	-0.0041	-0.0050	-0.0050	-0.0051	-0.0055	-0.0057	-0.0056	-0.0053	-0.0055	-0.0064	-0.0094	-0.0033
R^2	0.25	0.20	0.19	0.19	0.18	0.16	0.13	0.11	0.11	0.15	0.28	0.07
E[R]-1	1.89	1.82	1.85	1.90	1.96	2.02	2.10	2.46	4.89	18.70	97.08	271.19
E[$R^b - 1$]	0.13	0.99	0.96	0.96	0.95	0.95	0.94	0.93	0.86	0.48	-0.59	1.11
UBS Survey Data:												
Corr[PD$_t$,$E_t^{\mathcal{P}}$R$_{t,t+4}$]	0.79	0.79	0.80	0.81	0.84	0.85	0.87	0.88	0.91	0.90	0.92	0.96
Percent of periods w												
multiple equilibria:		0.00	0.00	0.27	0.37	0.41	0.56	1.31	5.60	34.49	78.77	97.13
Expected utility:												
Lagged updating		-4.02	-4.02	-4.01	-4.01	-4.00	-3.99	-3.98	-3.97	-3.89	-3.39	-2.46
Current updating		-4.07	-4.06	-4.06	-4.06	-4.06	-4.06	-4.06	-4.06	-4.11	-4.24	-4.02

Table A7: Asset pricing moments with simultaneous belief updating
(market clearing price furthest away from previous period price)

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, see table 3 for a description of the labels used in the first column, and the asset pricing moments for the estimated model from table 3 (diagonal matrix), considering different values for the share α of belief updaters using current price growth observations (columns 3 to 13). The table also reports the percentage share of periods in which multiple equilibria are encountered along the equilibrium (fifth to last line) and the unconditional expected utility of agents using current or lagged price growth observations for updating beliefs (last two lines). In periods with multiple equilibria, the equilibrium price furthest away from the previous market clearing price is selected.