Appendix to Fuster, Laibson, and Mendel (2010)

This appendix is divided into four sections. In section A.1, we report the maximum likelihood estimates that were used to create the impulse response functions in Exhibit 2. In section A.2, we report results from Monte-Carlo analysis that suggest that ARIMA(0,1,q) with “large” q are better able to capture low-frequency mean reversion than lower-order ARIMA(p,1,q) with p and $q \leq 3$. Moreover, we show that large q estimates are not subject to a substantial bias in the predicted persistence of the series when the true data generating process is not nested by the ARIMA(0,1,q) structure. In section A.3, we provide a more detailed description of the structural model in the paper. In section A.4, we report quantitative simulation results from the model and compare them with actual empirical moments.

A.1: Maximum Likelihood Estimates Underlying Exhibits 2a-2d

In this section, we report the maximum likelihood estimates that were used to create the impulse response functions in Exhibits 2a-2d. We show estimates for four different models: ARIMA(1,1,0), ARIMA(0,1, q_1), ARIMA(0,1, q_2), ARIMA(0,1, q_3). Here the index on q represents the number of years that are covered by the moving average (MA) terms. For quarterly data, $q_1 = 4,\ q_2 = 8,\ q_3 = 12.$ For monthly data, we estimate the analogous values: $q_1 = 12,\ q_2 = 24,\ q_3 = 36.$ Note that Exhibits 2a-2d only plot the impulse response functions for the cases ARIMA(1,1,0) and ARIMA(0,1, q_3).

The last line in the tables reports the predicted long-term persistence of a one-unit shock that is implied by the ARMA coefficients. This persistence is given by (1+sum of MA coefficients)/(1-sum of AR coefficients).

Table "2a": Real GDP (quarterly data)

	Log nominal GDP (x_t)							
	ARIMA (1,1,0)		ARIMA (0,1,4)		ARIMA (0,1,8)		ARIMA (0,1,12)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
x_{t-1}	0.363	(0.051)						
ε_{t-1}			0.337	(0.054)	0.326	(0.057)	0.333	(0.058)
ε_{t-2}			0.259	(0.057)	0.245	(0.057)	0.251	(0.062)
ε_{t-3}			0.079	(0.066)	0.066	(0.066)	0.045	(0.076)
ε_{t-4}					-0.048	(0.062)	-0.020	(0.064)
ε_{t-5}					-0.161	(0.064)	-0.147	(0.066)
ε_{t-6}					-0.045	(0.063)	-0.074	(0.066)
ε_{t-7}					-0.120	(0.072)	-0.120	(0.079)
ε_{t-8}					-0.133	(0.059)	-0.099	(0.075)
ε_{t-9}							-0.016	(0.074)
ε_{t-10}							0.053	(0.079)
ε_{t-11}							0.102	(0.065)
ε_{t-12}							-0.058	(0.060)
Constant	0.008	(0.001)	0.008	(0.001)	0.008	(0.001)	0.008	(0.001)
σ_ε	0.009	(0.000)	0.009	(0.000)	0.009	(0.000)	0.009	(0.000)
AIC	-1632.2		-1631.4		-1631.3		-1628.0	
BIC	-1621.7		-1610.2		-1596.1		-1578.6	
# obs.	251		251		251		251	
persistence	1.57		1.67		1.13		1.25	

Table “2b”: Unemployment (monthly data)

	Unemployment (x_t)							
	ARIMA (1,1,0)		ARIMA (0,1,12)		ARIMA (0,1,24)		ARIMA (0,1,36)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
x_{t-1}	0.120	(0.022)						
ε_{t-1}			-0.012	(0.028)	-0.032	(0.029)	-0.029	(0.032)
ε_{t-2}			0.199	(0.034)	0.197	(0.037)	0.221	(0.038)
ε_{t-3}			0.173	(0.036)	0.172	(0.036)	0.164	(0.038)
ε_{t-4}			0.172	(0.039)	0.174	(0.041)	0.159	(0.042)
ε_{t-5}			0.230	(0.034)	0.227	(0.037)	0.230	(0.038)
ε_{t-6}			0.134	(0.033)	0.118	(0.035)	0.120	(0.038)
ε_{t-7}			0.170	(0.037)	0.111	(0.039)	0.103	(0.040)
ε_{t-8}			0.175	(0.039)	0.105	(0.042)	0.096	(0.042)
ε_{t-9}			0.140	(0.034)	0.084	(0.037)	0.075	(0.038)
ε_{t-10}			0.023	(0.036)	-0.015	(0.038)	-0.016	(0.040)
ε_{t-11}			0.070	(0.036)	0.038	(0.038)	0.051	(0.038)
ε_{t-12}			-0.197	(0.036)	-0.228	(0.041)	-0.249	(0.042)
ε_{t-13}					-0.040	(0.039)	-0.069	(0.042)
ε_{t-14}					-0.104	(0.036)	-0.100	(0.038)
ε_{t-15}					-0.049	(0.036)	-0.043	(0.042)
ε_{t-16}					-0.056	(0.042)	-0.069	(0.041)
ε_{t-17}					-0.071	(0.041)	-0.094	(0.039)
ε_{t-18}					-0.027	(0.040)	-0.046	(0.043)
ε_{t-19}					0.012	(0.044)	-0.017	(0.046)
ε_{t-20}					0.015	(0.039)	-0.005	(0.042)
ε_{t-21}					-0.035	(0.037)	-0.058	(0.044)
ε_{t-22}					0.055	(0.040)	0.036	(0.041)
ε_{t-23}					-0.020	(0.038)	-0.069	(0.041)
ε_{t-24}					-0.185	(0.035)	-0.205	(0.041)
ε_{t-25}							-0.035	(0.043)
ε_{t-26}							-0.115	(0.043)
ε_{t-27}							0.005	(0.042)
ε_{t-28}							-0.072	(0.042)
ε_{t-29}							0.007	(0.045)
ε_{t-30}							-0.002	(0.040)
ε_{t-31}							-0.052	(0.043)
ε_{t-32}							-0.101	(0.040)
ε_{t-33}							-0.029	(0.040)
ε_{t-34}							0.088	(0.040)
ε_{t-35}							-0.049	(0.040)
ε_{t-36}							-0.049	(0.037)
Constant	0.008	(0.010)	0.009	(0.017)	0.007	(0.011)	0.005	(0.007)
σ_ε	0.214	(0.003)	0.194	(0.003)	0.189	(0.004)	0.186	(0.004)
AIC	-175.43		-297.71		-315.35		-311.65	
BIC	-161.58		-233.05		-195.26		-136.14	
# obs.	749		749		749		749	
persistence	1.14		2.28		1.44		0.78	

Table "2c": Real earnings (quarterly data)

	Real earnings (x_t)							
	ARIMA (1,1,0)		ARIMA (0,1,4)		ARIMA (0,1,8)		ARIMA (0,1,12)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
x_{t-1}	0.202	(0.051)						
ε_{t-1}			0.186	(0.053)	0.190	(0.057)	0.188	(0.061)
ε_{t-2}			0.027	(0.058)	0.035	(0.063)	-0.002	(0.064)
ε_{t-3}			0.131	(0.057)	0.092	(0.063)	0.064	(0.060)
ε_{t-4}			-0.134	(0.054)	-0.165	(0.060)	-0.162	(0.061)
ε_{t-5}					-0.204	(0.066)	-0.200	(0.069)
ε_{t-6}					-0.227	(0.062)	-0.185	(0.067)
ε_{t-7}					-0.094	(0.066)	-0.001	(0.065)
ε_{t-8}					-0.223	(0.060)	-0.180	(0.059)
ε_{t-9}							0.036	(0.070)
ε_{t-10}							-0.001	(0.068)
ε_{t-11}							-0.004	(0.075)
ε_{t-12}							-0.199	(0.067)
Constant	0.008	(0.002)	0.008	(0.002)	0.008	(0.001)	0.008	(0.001)
σ_ε	0.025	(0.001)	0.025	(0.001)	0.024	(0.001)	0.023	(0.001)
AIC	-1136.2		-1136.6		-1147.2		-1147.1	
BIC	-1125.6		-1115.4		-1111.9		-1097.7	
# obs.	251		251		251		251	
persistence	1.25		1.21		0.41		0.36	

Table “2d”: S&P 500 (monthly data)

	Excess Returns (x_t)							
	ARIMA (1,1,0)		ARIMA (0,1,12)		ARIMA (0,1,24)		ARIMA (0,1,36)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
x_{t-1}	0.114	(0.018)						
ε_{t-1}			0.121	(0.021)	0.111	(0.024)	0.113	(0.027)
ε_{t-2}			0.006	(0.025)	-0.009	(0.026)	-0.008	(0.027)
ε_{t-3}			-0.104	(0.025)	-0.090	(0.025)	-0.088	(0.026)
ε_{t-4}			0.020	(0.026)	0.023	(0.029)	0.033	(0.030)
ε_{t-5}			0.082	(0.026)	0.078	(0.029)	0.077	(0.030)
ε_{t-6}			-0.024	(0.024)	-0.033	(0.027)	-0.033	(0.028)
ε_{t-7}			0.017	(0.027)	-0.006	(0.028)	-0.008	(0.029)
ε_{t-8}			0.026	(0.020)	0.032	(0.024)	0.030	(0.025)
ε_{t-9}			0.083	(0.020)	0.069	(0.025)	0.068	(0.026)
ε_{t-10}			0.000	(0.024)	0.020	(0.026)	0.010	(0.027)
ε_{t-11}			-0.010	(0.024)	-0.029	(0.027)	-0.023	(0.028)
ε_{t-12}			0.028	(0.027)	0.015	(0.027)	0.015	(0.029)
ε_{t-13}					-0.058	(0.028)	-0.041	(0.031)
ε_{t-14}					-0.084	(0.029)	-0.084	(0.029)
ε_{t-15}					0.024	(0.027)	0.014	(0.028)
ε_{t-16}					-0.030	(0.029)	-0.040	(0.030)
ε_{t-17}					0.096	(0.027)	0.094	(0.029)
ε_{t-18}					0.023	(0.030)	0.025	(0.033)
ε_{t-19}					-0.051	(0.027)	-0.049	(0.029)
ε_{t-20}					-0.112	(0.031)	-0.117	(0.033)
ε_{t-21}					-0.133	(0.033)	-0.135	(0.034)
ε_{t-22}					-0.030	(0.031)	-0.018	(0.033)
ε_{t-23}					-0.055	(0.028)	-0.052	(0.031)
ε_{t-24}					0.041	(0.029)	0.031	(0.031)
ε_{t-25}							-0.051	(0.032)
ε_{t-26}							0.032	(0.031)
ε_{t-27}							-0.003	(0.031)
ε_{t-28}							0.009	(0.035)
ε_{t-29}							-0.021	(0.033)
ε_{t-30}							0.008	(0.035)
ε_{t-31}							0.034	(0.034)
ε_{t-32}							-0.015	(0.035)
ε_{t-33}							-0.018	(0.033)
ε_{t-34}							-0.046	(0.033)
ε_{t-35}							0.043	(0.034)
ε_{t-36}							0.002	(0.035)
Constant	0.006	(0.002)	0.006	(0.002)	0.006	(0.002)	0.006	(0.002)
σ_ε	0.054	(0.054)	0.053	(0.053)	0.052	(0.052)	0.052	(0.052)
AIC	-3007.13		-3010.46		-3025.76		-3010.01	
BIC	-2992.40		-2941.72		-2898.10		-2823.44	
# obs.	1002		1002		1002		1002	
persistence	1.13		1.25		0.81		0.79	

A.2 Monte-Carlo Analysis

In this section, we report results from a Monte-Carlo analysis that suggests that ARIMA(0,1,q) with "large" q (e.g., four years of lags) are better able to capture empirically relevant low-frequency mean reversion than lower-order ARIMA(p,1, q) with p and $q \leq 3$. In addition, we show that ARIMA(0,1, q) models with large q are not subject to substantial bias in the predicted persistence of the series when the true data generating process is *not* nested by the ARIMA(0,1, q) structure.

We focus our analysis on the natural logarithm of the real U.S. net operating surplus of private enterprises. This variable, which is a proxy for capital income, is the driving process in the model in the paper.

Our analysis is conducted as follows:

We start by estimating a total of 8 statistical models on the available data:

- ARIMA(1,0,1), ARIMA(2,0,2), ARIMA(3,0,3); using linearly detrended data
- ARIMA(1,1,1), ARIMA(2,1,2), ARIMA(3,1,3), ARIMA(0,1,12), ARIMA(0,1,16).

The first group of models is stationary and the second group is non-stationary.

Then, we use the estimated models as the true data generating process (DGP) to simulate 200 samples of 252 periods; this simulated sample length matches the sample length of the empirical data. Hence we have eight models, each of which has 200 simulated samples, or 1600 simulated samples in total. For each of these 1600 simulated samples, we estimate the following five models: ARIMA(0,1,4), ARIMA(0,1,8), ARIMA(0,1,12), ARIMA(2,1,2), and ARIMA(3,1,3). Hence, we are generating 8000 maximum likelihood estimations. As expected, about 5% of the estimates are automatically discarded by Stata (the statistical software we use) because of a failure in the convergence algorithm.

One well-known problem when estimating moving average (MA) models is that the maximum likelihood estimator of the moving average root (i.e., the negative of the sum of the lagged MA coefficients) tends to "pile up" at a value of 1, which means that the predicted long-term persistence of a shock equals 0. This pileup occurs because the sample likelihood function is locally flat at an MA root of 1, so that it is a local maximum of the likelihood function and may be the global maximum in finite samples, even if the true MA root is less than unity (see Campbell and Mankiw 1987 or Stock 1994 for discussions).

As anticipated, in our simulations we observe that the estimated MA root is often *exactly* equal to 1. In the case of stationary DGPs with an *infinite* sample, this is of course what the estimation "should" find, because the true persistence of a shock equals zero. For the non-stationary DGPs that we use, however, true persistence is never zero. There is no simple solution for dealing with estimation results that imply exactly zero persistence, because it is impossible to know with certainty whether the likelihood function actually does have a global maximum away from the unit MA root.

To deal with this issue, we report in each case the percentage of simulations for which the MA root was exactly unity (or more precisely, the absolute value of the sum of MA coefficients was within 0.0001 of -1) and the mean estimated long-term persistence including or dropping these cases.

We furthermore report the percentage of simulations for which the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) -- also called the Schwarz Criterion -- would select either the ARIMA(2,1,2) or the ARIMA(3,1,3) over the ARIMA(0,1,q) models.

Stationary data generating processes

In all of the following three cases, the true model that is used to generate simulated data is stationary, and consequently the true underlying persistence is 0.

Case 1: Data generating process is ARIMA(1,0,1)

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	0.99	0.63	0.35	0.35	0.28
Std(est. persistence)	0.18	0.30	0.35	0.49	0.49
% of simulations with exact zero	0.0%	8.4%	36.1%	59.2%	66.0%
Mean est. pers. w/o exact zeros	0.99	0.68	0.55	0.86	0.83

The AIC and BIC select either the ARIMA(2,1,2) or (3,1,3) in 90.1% and 88.5% of all simulations, respectively.

Case 2: Data generating process is ARIMA(2,0,2)

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	1.20	0.56	0.22	0.67	0.29
Std(est. persistence)	0.19	0.36	0.33	0.62	0.52
% of simulations with exact zero	0.0%	16.1%	57.3%	40.6%	69.3%
Mean est. pers. w/o exact zeros	1.20	0.67	0.51	1.12	0.96

The AIC and BIC select either the ARIMA(2,1,2) or (3,1,3) in 89.6% and 85.9% of all simulations, respectively.

Case 3: Data generating process is ARIMA(3,0,3)

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	1.36	0.50	0.20	0.75	0.91
Std(est. persistence)	0.21	0.37	0.31	0.66	0.61
% of simulations with exact zero	0.0%	21.9%	57.1%	39.3%	26.5%
Mean est. pers. w/o exact zeros	1.36	0.65	0.47	1.23	1.24

The AIC and BIC select either the ARIMA(2,1,2) or ARIMA(3,1,3) in 50.5% and 64.3% of all simulations, respectively.

We conclude that all estimated processes fail to consistently detect the mean-reversion in the true data generating process. This problem is particularly pronounced for the ARIMA(2,1,2) and ARIMA(3,1,3) models when the data generating process becomes relatively complicated (Case 3). The ARIMA(0,1,12) models tends to predict low persistence; it often estimates an MA root of unity but even if it doesn't, it on average predicts a persistence well below 1, unlike the low-order models. Despite the superior persistence predictions of the ARIMA(0,1,12), the information criteria usually select either the ARIMA(2,1,2) or ARIMA(3,1,3).

Non-stationary data generating processes

In the following five cases, the true model that is used to generate simulated data is non-stationary.

Case 4: Data generating process is ARIMA(1,1,1) with true persistence of 1.25

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	1.25	1.21	1.20	1.10	1.09
Std(est. persistence)	0.15	0.23	0.32	0.48	0.54
% of simulations with exact zero	0.0%	0.5%	1.1%	11.6%	13.8%
Mean est. pers. w/o exact zeros	1.25	1.22	1.21	1.24	1.26

The AIC and BIC select either the ARIMA(2,1,2) or ARIMA(3,1,3) in 85.0% and 75.6% of all simulations, respectively.

Case 5: Data generating process is ARIMA(2,1,2) with true persistence of 1.29

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	1.24	1.22	1.22	1.19	1.11
Std(est. persistence)	0.14	0.24	0.33	0.53	0.56
% of simulations with exact zero	0.0%	2.6%	10.9%	32.8%	41.7%
Mean est. pers. w/o exact zeros	1.24	1.22	1.23	1.34	1.27

The AIC and BIC select either the ARIMA(2,1,2) or ARIMA(3,1,3) in 90.1% and 82.3% of all simulations, respectively.

Case 6: Data generating process is ARIMA(3,1,3) with true persistence of 1.33

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	1.33	1.24	1.19	1.10	1.31
Std(est. persistence)	0.15	0.24	0.34	0.57	0.60
% of simulations with exact zero	0.0%	0.5%	1.6%	12.5%	5.7%
Mean est. pers. w/o exact zeros	1.33	1.25	1.21	1.26	1.39

The AIC and BIC select either the ARIMA(2,1,2) or (3,1,3) in 80.2% and 71.9% of all simulations, respectively.

Case 7: DGP is ARIMA(0,1,12) with true persistence of 0.36

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	1.23	0.29	0.20	0.87	0.78
Std(est. persistence)	0.27	0.24	0.23	0.62	0.60
% of simulations with exact zero	0.0%	30.2%	48.7%	24.1%	21.1%
Mean est. pers. w/o exact zeros	1.23	0.42	0.38	1.14	0.99

The AIC and BIC select either the ARIMA(2,1,2) or (3,1,3) in 7.0% and 67.3% of all simulations, respectively. (The AIC would select the ARIMA(2,1,2) or ARIMA(3,1,3) over the ARIMA(0,1,8) in 23.6% of cases.)

Case 8: DGP is ARIMA(0,1,16) with true persistence of 0.29

	(0,1,4)	(0,1,8)	(0,1,12)	(2,1,2)	(3,1,3)
Mean estimated persistence	1.25	0.28	0.22	0.88	1.00
Std(est. persistence)	0.26	0.26	0.30	0.62	0.52
% of simulations with exact zero	0.0%	35.4%	52.5%	27.3%	15.7%
Mean est. pers. w/o exact zeros	1.25	0.43	0.46	1.22	1.18

The AIC and BIC select either the ARIMA(2,1,2) or (3,1,3) in 6.6% and 64.6% of all simulations, respectively. (The AIC would select the ARIMA(2,1,2) or (3,1,3) over the ARIMA(0,1,8) in 24.2% of cases.)

We conclude that in cases 4, 5, and 6, where the true data generating process is characterized by persistence above 1, the persistence implied by the estimated ARIMA(0,1,12) model is somewhat downward biased, but not severely so. In cases 7 and 8, on the other hand, where the true data generating process is characterized by persistence well below 1, the ARIMA(0,1,8) and ARIMA(0,1,12) models are the only ones that on average come close to estimating the right persistence, while lower-order models usually fail to predict any mean reversion. Nevertheless, the information criteria frequently choose one of the low-order models (particularly the BIC).

The pileup problem discussed above is severe for the ARIMA(0,1,12) model in cases 7 and 8 and drives the slight downward bias apparent in the first lines in the respective tables. If runs where this occurs are disregarded, the estimated persistence is slightly upward-biased. In the impulse response functions in Exhibits 2a-2d of the paper, implied persistence never equals zero.

Overall, we believe that the slight downward bias we find for the ARIMA(0,1,q) for large q is outweighed by this specification's ability to detect mean reversion even in samples of the size we use in our empirical work. Low-order models, on the other hand, systematically over-predict persistence when the true underlying process is partially mean reverting.

A.3 Detailed description of model

Consider a representative agent economy with two Lucas-style trees. An equity tree generates stochastic dividends, x_t, and a labor tree generates deterministic labor income, y. Dividends can be decomposed into a fixed component, μ, and a zero-mean stochastic process given by

$$x_{t+1} = \alpha x_t + \beta x_{t-1} + \eta_{t+1} \tag{1}$$

Agents hold intuitions generated by

$$\Delta x_{t+1} = \phi \Delta x_t + \varepsilon_{t+1}. \tag{2}$$

For example, equation (2) is the model that is almost always chosen by the Bayesian Information Criterion from the ARIMA(p,1,0) class when the true data-generating process is equation (1).[1] Assume that parameter ϕ is pinned down using ordinary least

[1] This assumes that the econometrician has a relevant macro sample of 252 quarters of data.

squares[2], so that $\phi = \frac{\alpha - \beta - 1}{2}$. Hence, ϕ is not a free parameter. This simple growth regression is used to form intuitive forecasts, which we express as $I_t\left[x_{t+\tau}\right]$.

Agents form beliefs – which we call natural expectations -- that are a weighted combination of intuitive expectations and rational expectations.

$$N_t\left[x_{t+\tau}\right] \equiv \lambda I_t\left[x_{t+\tau}\right] + \left(1-\lambda\right) E_t\left[x_{t+\tau}\right] \tag{3}$$

The representative agent has time-separable quadratic preferences, and discount factor δ. We study an open economy with foreign lenders who are willing to borrow and lend at a constant risk-free (gross) rate of interest R. To enhance tractability, we assume that $\delta R = 1$, which implies that consumption is proportional to the discounted value of forecasted claims, where the discount factor is $1/R$.

The economy's dynamic budget constraint is given by,

$$B_{t+1} = c_t + RB_t - x_t - y,$$

where B_t is the debt to foreign agents in period *t*, and c_t is consumption in period *t*. For simplicity, we assume that foreign economic agents have no claims on domestic capital. This is consistent with non-rational home bias or rational non-diversification coming

---

[2] To derive the value of ϕ, start with the true data generating process for *x* and difference it to obtain:

$$\Delta x_t = \alpha \Delta x_{t-1} + \beta \Delta x_{t-2} + \varepsilon_t - \varepsilon_{t-1}. \tag{1}$$

Now multiply the RHS and LHS of (1) by $\Delta x_{t-\tau}$ and take expectations. Specifically, execute this step three times, for $\tau = 0, 1, 2$. These three calculations will yield a system of second moments:

$$V(\Delta x) = \alpha Cov(\Delta x, \Delta x_{-1}) + \beta Cov(\Delta x, \Delta x_{-2}) + (2-\alpha)\sigma_\varepsilon^{\ 2}$$
$$Cov(\Delta x, \Delta x_{-1}) = \alpha V(\Delta x) + \beta Cov(\Delta x, \Delta x_{-1}) - \sigma_\varepsilon^{\ 2}$$
$$Cov(\Delta x, \Delta x_{-2}) = \alpha Cov(\Delta x, \Delta x_{-1}) + \beta V(\Delta x)$$

Solve this system to recover the value of ϕ, which is an OLS coefficient:

$$\phi = \frac{Cov(\Delta x, \Delta x_{-1})}{V(\Delta x)} = \frac{\alpha - \beta - 1}{2}.$$

through an agency channel.[3] Our results would be qualitatively similar as long as domestic agents hold a disproportionate share of the domestic capital stock.

We simulate the model at quarterly frequencies. The quarterly risk-free gross return is $R = 1.015$. The parameters on the true data generating process—equation **(1)**—are $\alpha = 1.16$ and $\beta = -0.24$. These values are estimated by first detrending the constant dollar NIPA data for the net operating surplus of private businesses[4] and then estimating equation **(1)**. This procedure is biased because it imposes stationarity. We study this extreme case to illustrate ideas rather than because we believe that there is no persistent component in capital income. The stationarity property can be relaxed without meaningfully changing the results of this paper. We set the standard deviation of innovations in equation **(1)** to $\sigma = 0.05$. This is done so that equilibrium (annual) asset volatility matches the historical average standard deviation of 0.17. We set the average capital income share, $\mu / (\mu + y)$, to 1⁄3, which matches the historical average. Finally, we take a shot in the dark and set the intuitive forecasting weighting parameter in equation **(3)** to $\lambda =$ ½. In other words, intuitive forecasting and rational forecasting are equally influential in the mind of the representative agent. In future work, this parameter could be estimated using a moment-based method. Finally, to enhance the simplicity of our closed form expressions, we let risk aversion become vanishingly small, so that asset pricing can be done with the risk free rate. Specifically, we study the limiting economy that emerges as the weight on the quadratic term in the utility function goes to zero.

Under these assumptions, the intuitive forecast for discounted equity returns is,

$$I_t\left[\sum_{s=1}^{\infty}\frac{x_{t+s}}{R^s}\right] = \frac{\mu + x_t}{R-1} + \Delta x_t \frac{\phi}{1-\phi}\left(\frac{\frac{1}{R}}{1-\frac{1}{R}} - \frac{\frac{\phi}{R}}{1-\frac{\phi}{R}}\right).$$

The rational forecast for discounted equity returns is,

---

[3] For example, moral hazard and adverse selection make it inefficient for foreign investors to own most of the U.S. residential housing stock.

[4] NIPA accounts, 1947:1 to 2009:4, Table 1.10, line 12. See discussion in previous section.

$$E_t\left[\sum_{s=1}^{\infty}\frac{x_{t+s}}{R^s}\right]=\frac{\mu}{R-1}+A_t\left(\frac{\frac{r_1}{R}}{1-\frac{r_1}{R}}\right)+B_t\left(\frac{\frac{r_2}{R}}{1-\frac{r_2}{R}}\right)$$

where, $r_1=\frac{\alpha+\sqrt{\alpha^2+4\beta}}{2}$, $r_2=\frac{\alpha-\sqrt{\alpha^2+4\beta}}{2}$, $A_t = r_1\frac{x_t-r_2x_{t-1}}{r_1-r_2}$, and $B_t = r_2\frac{r_1x_{t-1}-x_t}{r_1-r_2}$.

For the purposes of Exhibit 3a, we define the cumulative *t*-period gross return as

$$R_{1,t}=\frac{\left(\sum_{i=1}^{t}d_iR^{t-i}\right)+p_t}{p_0}.$$

This definition implicitly assumes that equity dividends are reinvested in the *safe* asset with gross annual return R. The cumulative *excess* return is then

$$\pi_{1,t}=\frac{R_{1,t}}{R^t}-1.$$

These definitions have the desirable property that the asymptotic cumulative return for the Natural Expectations model matches the asymptotic cumulative return for the Rational Expectations model. Intuitively, agents realize the same dividend payouts from the Lucas tree whether or not they have rational expectations; the cumulative excess return in the long-run is therefore unaffected by the nature of belief formation.

Finally, we have also analyzed an extended version of the model with slow adjustment in consumption, which is like a consumption habit. In this "habit" version of the model, we assume that consumption is an equal-weighted average of last period's consumption and "target" consumption. Target consumption is defined as the annuity value of wealth (this is the consumption level implied by the model without habits). Adding a habit improves the model's empirical fit with respect to high frequency consumption dynamics. Without the habit, the model predicts that consumption responds instantly to wealth innovations, which leads to counterfactually high consumption volatility. The habit reduces this volatility, without materially affecting the predictability properties we are focusing on (reported in Table A.1, below).

A.4 Empirical evaluation

The following table reports the degree of return predictability using both empirical moments ("data") and simulated moments ("model"). The moments match at almost all horizons. For the simulations, we use the calibrated model to generate 1000 samples, each of which has 252 periods. Hence, our real and simulated samples have matched lengths.

Predictability of excess returns

τ	$corr\left(\frac{p_t}{e_t^{40}}, r_{t+\tau}\right)$		$corr\left(r_t, r_{t+\tau}\right)$		$corr\left(\Delta c_t, r_{t+\tau}\right)$	
	DATA	MODEL	DATA	MODEL	DATA	MODEL
1	-0.12	-0.10	0.09	0.00	-0.03	0.00
2	-0.14	-0.11	-0.05	-0.04	0.00	-0.04
3	-0.13	-0.10	-0.03	-0.05	-0.06	-0.05
4	-0.14	-0.09	-0.02	-0.05	-0.08	-0.05
5	-0.14	-0.08	-0.05	-0.04	-0.06	-0.04
6	-0.12	-0.07	-0.06	-0.04	-0.06	-0.04
7	-0.12	-0.07	-0.13	-0.03	-0.02	-0.03
8	-0.10	-0.06	0.01	-0.03	-0.08	-0.03

Caption: Correlations calculated with U.S. data (1947:1 to 2010:2) and simulated data The ratio p_t / e_t^{40} is the S&P 500 index divided by 40 quarters of inflation-adjusted earnings (Campbell and Shiller 1988). P/E ratios come from Robert Shiller's website. Excess returns come from Kenneth French's website: "the value-weighted return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate (from Ibbotson Associates)." Real consumption is from the NIPA accounts. For the data we use changes in log consumption. For the model we use changes in the absolute level of consumption, since the simulated model has no long-run growth.

References

Campbell, John Y., and N. Gregory Mankiw. 1987. "Are Output Fluctuations Transitory?" *The Quarterly Journal of Economics*, 102(4): 857-880.

Campbell, John Y., and Robert J. Shiller. 1988. "Stock Prices, Earnings, and Expected Dividends." *The Journal of Finance*, 43(3): 661-676.

Stock, James H., 1994. "Unit Roots, Structural Breaks and Trends." in: *Handbook of Econometrics*, Volume IV, eds. R.F. Engle and D.L. McFadden.