Return predictability and social dynamics

Richard Hule‡ Jochen Lawrenz§

Nov 2009

‡Richard Hule, Department of Economics, Innsbruck University, A-6020 Innsbruck, Austria, richard.hule@uibk.ac.at, Phone +43-512-507-7333.
§Dr. Jochen Lawrenz, Department of Banking & Finance, Innsbruck University, A-6020 Innsbruck, Austria, jochen.lawrenz@uibk.ac.at, Phone +43-512-507-7582.
Return predictability and social dynamics

Nov 2009

Abstract

The possibility to predict stock returns from financial ratios is a long-standing but still controversial topic. There is ongoing debate about the empirical evidence as well as about appropriate theoretical explanations. We provide evidence from a simulated economy that local, social interaction among agents is remarkably successful in matching several empirically established facts. We find significant return predictability at various forecast horizons, absence of dividend growth predictability, high persistence in dividend yields and absence of significant return autocorrelations. Our results suggest that social dynamics are a simple, intuitively appealing and successful way to explain predictability.

JEL classification: G17, G12, D80, E37

Keywords: Return predictability, Predictive regressions, Social dynamics, Hidden Markov chain, Agent-based models
1 Introduction

The predictability of asset returns is a long-standing and controversial topic in empirical as well as in theoretical financial economics. While on the empirical side, the evidence for predictability is still not a settled issue,\(^1\) the theoretical literature has already put forward different attempts how to reconcile predictability with rational investor behavior. In this paper, we contribute to the latter strand of the literature by showing that social interaction within an agent-based model is surprisingly successful in matching many of the stylized facts that have been established in various empirical contributions. The crucial point where our approach differs from previous work is that we drop the representative agent assumption. We consider an economy populated by a finite number of risk-averse agents who make consumption decisions conditional on a set of influencing factors. Among others, we assume that agents are able to observe past decisions of a subset of other agents, say, their neighbors. Depending on the weight an agent puts on the decisions of her neighborhood, persistent clusters of opinion can emerge. Our results show that a high propensity to interact locally gives rise to persistence in the predictive variable, which is known to yield evidence for predictability.\(^2\)

The basic workhorse for much research in financial economics is a consumption-based asset pricing model within a pure-exchange economy, where consumption decisions are analyzed by assuming a representative agent with constant risk preferences. Such a basic setup cannot account for predictability. It is straightforward to show that such a model predicts the price-dividend ratio, which is most frequently used as predictive variable,\(^3\) to be constant. Therefore, more sophisticated models are needed to match the observable time-variation in the price-dividend ratio as well as its ability to forecast future returns. Previous asset pricing models that have been put forward to explain predictability can roughly be classified into two broad categories. On the one hand, there are models which retain the assumption of a simple, standard endowment process, but assume a more sophisticated representative agent. The additional complexity may come

---

\(^1\) See e.g. Cochrane (2008), Lettau & Van Nieuwerburgh (2008), or Koijen & Van Binsbergen (2009) for recent contributions that support the evidence of predictability, while Ang & Bekaert (2007), or Goyal & Welch (2003, 2008) are examples that question the evidence.

\(^2\) Collard et al. (2006) stress the role of persistence in explaining predictability.

\(^3\) See e.g. Campbell & Shiller (1989) for an early study.
from time-varying risk aversion, habit formation, long-run consumption risk or time-varying risk-sharing opportunities.\textsuperscript{4} Loosely speaking, this way the simple endowment process is transformed to a price process which has more temporal structure and may display predictability of returns. On the other hand, models have been put forward, which retain the assumption of a simple, standard representative agent, but which assume that the underlying process for the state variable already exhibits a more complex structure.\textsuperscript{5} This way, an already complex endowment process feeds into a simple agent and yields the predictability evidence. Both approaches stay within the representative agent paradigm. We offer a third, alternative way of thinking about predictability. We assume both, a simple endowment process \textit{as well as} simple agents, and show that introducing a simple form of local interaction (or social dynamics) among agents is able to produce time series that exhibit many of the observable stylized facts of predictability.

Our model combines three main features that can be considered as important, salient characteristics of financial markets. First, we assume that the growth rate of the economy can switch between two states. Second, agents have incomplete information with respect to the state of the economy. Third, we consider agents to interact locally, thereby inducing social dynamics. The fact that the growth rate can switch between two states captures the evidence that the economy changes between times of expansion and times of contractions, and therefore reflects business cycles. Arguably, Hamilton (1989) was one of the first contribution to apply a regime-switching model to economic business cycles. More recently, in particular Cecchetti et al. (1990, 2000), Whitelaw (2000), or Bansal & Yaron (2004) use regime-switching models to analyze various aspects of asset returns.\textsuperscript{6} While the issue of predictability is also mentioned in Cecchetti et al. (2000), in particular Lettau & Van Nieuwerburgh (2008) argue that tak-

\begin{itemize}
  \item[4] See e.g. Menzly et al. (2004) for time-varying preference, Campbell & Cochrane (1999) for habit formation. For an overview, see e.g Cochrane (2005), or Koijen & Van Nieuwerburgh (2009).
  \item[5] Examples include Calvet & Fisher (2007), who consider a dividend process with state-dependent volatility or Bansal & Yaron (2004) who assume a dividend process whose growth rate follows an AR(1) process.
  \item[6] Cecchetti et al. (2000) address the well-known equity premium puzzle, while Whitelaw (2000) shows that regime-switching models can explain the non-linear relationship between expected returns and volatility.
\end{itemize}
ing account of regime-switches substantially strengthens the evidence for return predictability. Further contributions which feature state-dependent growth rates include Veronesi (2000), Brennan & Xia (2001), and Ozoguz (2009). In line with these latter works, our model shares the second main assumption that agents are assumed not to be able to observe the current state (i.e., growth rate) of the economy. Agents have incomplete information about the hidden Markov chain driving the growth rate of the endowment process. However, they can draw inference from the realizations of the endowment process by rational Bayesian learning principles. The best estimate of the state is then found by applying results from filtering theory.\textsuperscript{7} Veronesi (2000) and Brennan & Xia (2001) study the impact of incomplete information on the properties of stock returns, in particular on the stock price volatility and the equity premium. Brennan & Xia (2001) find that their model is able to explain excess volatility as well as a high equity premium. Veronesi (2000) on the other hand finds that the incomplete information makes the equity premium puzzle even more puzzling, since he shows that a lower precision of signals (i.e., worse information quality) tends to decrease the risk premium, rather than increasing it. Together they underscore the importance of taking into account the effects of incomplete information on equilibrium asset prices. More closely related to the predictability literature, Timmermann (1993) shows that learning can account for predictability.\textsuperscript{8} More recently, Ozoguz (2009) examines how incomplete information affects the pricing of risks. She estimates a Markov regime-switching model, and shows that a measure of investors’ beliefs and uncertainty can act as a priced state variable and has some ability to forecast future returns. However, in her model she sticks to the representative agent assumption. Therefore, time-variation in the investors’ beliefs about the state of the economy are exclusively due to the arrival of news. A similar setup is used by Cecchetti et al. (2000). Their approach differs in that they assume that the representative agent can observe the underlying Markov process, but that her beliefs about the transition probabilities are distorted. Thereby, they account for over-optimism in expansionary times and over-pessimism in contraction times. Our model complements the results of these contributions by dropping the representative agent assumption.

\textsuperscript{7} Our approach is similar to early work by Detemple (1986) or Dothan & Feldman (1986).

\textsuperscript{8} Predictability can be traced back to the persistence in the updated estimates. However, he considers the case where a representative agent has incomplete information with respect to constant drift and diffusion parameters. In such a setup, the predictability effect fades out quickly, since a constant can be estimated with high accuracy after only few observations.
assumption. By assuming a finite number of agents, our approach allows the change in aggregate beliefs to be either the result of new global information, or the effect of socially interacting agents. Note that our model is able to replicate the results mentioned above, when we restrict each agent to act exactly the same way. Therefore, our results nest the findings from the representative agent literature.

This brings us to the third, and arguably most significant feature of our model: Social dynamics, or local interaction, in an agent-based economy. The bulk of research in neoclassical research relies on the assumption that aggregate decisions can be analyzed as if they were made by a single, representative agent. By construction, this abstracts away from any social interactions. Within the traditional asset pricing literature, the consideration of social dynamics has not received much attention. Exceptions are from the behavioral finance literature. A prominent proponent of the considerations of psychological and social aspects is Shiller, who made an early case for the importance of social dynamics in Shiller (1984). Although on a descriptive level, he makes a powerful case that social dynamics are an important characteristic of financial markets. However, the analysis of social dynamics has received much more attention in the literature that has become known under the notion of agent-based models. Kirman (1992) is a prominent example for a critique on the representative agent assumption, and in Kirman (1993), he puts forward a model that replicates the dynamics of opinions in a social context and explains apparent abrupt changes in aggregate behavior. As a more recent contribution, Lux (2009) presents empirical evidence that the formation of a business climate index is significantly affected by dynamics that are well described by local interaction and 'opinion contagion'. Our model is also close to Orlean (1995) and Cont & Bouchaud (2000). In Orlean (1995), agents update their beliefs not only on the basis of their private signal, but also by observing the collective group opinion. The model by Cont & Bouchaud (2000) has gained much recognition as a simple model that is able to produce

---

9 In the following, we will use the terms 'social dynamics' and 'local interaction' synonymously.

10 “Investing in speculative assets is a social activity. Investors spend a substantial part of their leisure time discussing investments, reading about investments, or gossiping about others’ successes or failures in investing. It is thus plausible that investors’ behavior (and hence prices of speculative assets) would be influenced by social movements.” Shiller (1984), p. 457.

11 Thereby, he generalizes the herding results to a non-sequential setting. See e.g. Bikchandani et al. (1992) for the classical sequential herding model.
stylized market facts that are consistent with empirical observations. They allow for a heterogeneous market structure, where different groups can make independent decisions. The formation of groups is assumed to be random, and the formal tool which they apply is random graph theory. We take up the basic intuition from these models to formalize the idea of social interaction. First, we choose a fixed network (see Wilhite, 2006), basically a cellular automaton,\textsuperscript{12} rather than endogenous formation of the links between the neighbors.\textsuperscript{13} Second, we do not assume that the state of each agent mechanically determines his demand, but rather that demand is the outcome of expected utility maximization subject to an agent’s beliefs. Third, in contrast to several models (e.g. Lux & Marchesi, 1999; Farmer & Joshi, 2002), we do not assume agents to act as chartists or trend followers, since in this way an automatic positive feedback mechanism is already model-built-in right from the outset.

Agent-based models have been shown to be successful in explaining a wide range of stylized market facts, such as volatility clustering, fat tails or scaling laws in returns and sudden shifts in aggregate behavior. Our contribution shows that social dynamics also help to explain more subtle financial market phenomena such as return predictability. Thus, in a broader sense, our contribution provides another step in the attempt to bridge the gap between the classical financial economics literature and the more recent surge of interest in agent-based models.

The remainder of the article is structured as follows. The next section describes the economy. In section 3, we discuss empirical evidence on predictability and put forward evidence from our simulated economy. Section 4 concludes.

2 The Economy

In line with a substantial part of the finance literature, we consider a pure-exchange economy in the spirit of Lucas (1978). There exists an exogenous stochastic non-storable endowment process which is traded in the form of one unit of stock, being infinitely divisible. We assume time to be discrete with an in-

\textsuperscript{12} See Wolfram (2002) on cellular automata.

\textsuperscript{13} This amounts to saying something about the relative speed of change of trading and social interactions. Namely, we assume that the social environment is stable and changes more slowly than the dividend process of the risky asset. (See e.g. Vriend, 2006).
finite horizon. At each point in time \( t \), the stock pays a dividend whose dynamics follow the discretized version of a standard diffusion process, i.e.

\[
D_t = D_{t-\Delta t} \exp \left\{ \left( \mu_t - \sigma_D^2 / 2 \right) \Delta t + \sigma_D \sqrt{\Delta t} \epsilon_t \right\}
\]

(1)

where the \( \epsilon_t \) are standard normal iid, and \( \sigma_D \) is constant.\(^{14}\)

While volatility is assumed to be time-invariant, we allow for time-variation in the drift term. We assume that \( \mu_t \) can take on two possible values \( \mu_t \in \{ \mu_h, \mu_l \} \) with \( \mu_h \geq \mu_l \). At any point in time, the growth rate is expected to switch between the two states with a given probability, which is summarized in the transition matrix,

\[
T = \begin{pmatrix}
1 - p_{h,l} & p_{h,l} \\
p_{l,h} & 1 - p_{l,h}
\end{pmatrix},
\]

where \( p_{h,l} \) is the transition probability that the growth rate jumps from \( \mu_h \) to \( \mu_l \) and correspondingly for \( p_{l,h} \). The dividend process can thus be characterized as a Markov switching process (also known as Hidden Markov Chain) with an underlying binomial state variable. Such a specification has been advanced by Hamilton (1989) to characterize the business cycle. Subsequently, in particular Cecchetti et al. (1990, 2000) applied the Markov switching model to an endowment process in a standard Lucas economy much like we do. Veronesi (2000) and Brennan & Xia (2001) consider similar processes in continuous-time where the drift term can assume a finite number \( n \) of states, or is assumed to follow an Ornstein-Uhlenbeck process.

Obviously, while all of the parameters of the model are assumed to be common knowledge, the hidden state is not observable. Although the true current growth rate of the economy cannot be observed, the observable realizations of the dividend stream can be used to draw inference of the current state. Since the process \( (D_{\tau})_{0 \leq \tau \leq t} \) is observable and can be characterized as a regime-switching model with Gaussian innovations, the inference problem can be solved by applying a variant of the Kalman filter.\(^{15}\) We denote the estimate (at time \( t \) conditional on the information set \( (D_{\tau})_{0 \leq \tau \leq t} \)) that the economy grows at the high rate, \( \mu_h \), as \( \hat{\mu}_t \),

\(^{14}\) In general, the agents opportunity set is frequently assumed to additionally contain a riskless asset which is in zero net supply. However, note that in this endowment economy, equilibrium requires that aggregate consumption equals aggregate dividends, so that all agents only hold the risky asset, and the rate of return on the riskless asset is only to be interpreted as the economy’s shadow riskfree interest rate.

\(^{15}\) See e.g. Hamilton (2008)
i.e. \( \hat{\mu}_t = \text{Prob}\{\mu_t = \mu_h | (\tau_t)_{t \leq t}\} \). Denote by \( n_{s,t} \), the density of the log dividend, \( d_t = \log D_t \) in regime \( s \) at time \( t \), which is normally distributed: \( d_t \sim \mathcal{N}(\mu_s, \sigma_D^2) \). Then for any \( t \), the updated probability \( \hat{\mu}_t \) follows from

\[
\hat{\mu}_t = \frac{n_{h,t} (p_\mu + \hat{\mu}_{t-1} - 2p_\mu \hat{\mu}_{t-1})}{n_{t,t} + (n_{h,t} - n_{t,t}) (p_\mu + \mu_{t-1} - 2p_\mu \hat{\mu}_{t-1})}.
\]

Note that \( \hat{\mu}_t \) is the estimated probability for being in regime \( h \), and not the estimate for \( \mu_t \) itself.

Note further that \( \hat{\mu}_t \) is to be interpreted as a rational estimate consistent with Bayesian updating. This is in line with e.g. Veronesi (2000), Brennan & Xia (2001) and others, but contrasts with e.g. Ceccheti et al. (2000), who assume that the representative agent has distorted beliefs. In particular, they consider the impact of biased beliefs about the transition probabilities (\( p_{h,t}, p_{t,h} \) in our notation) and show that this can reproduce some important facts about equity returns. In our setup, we consider all agents to have access to the information given by the unbiased estimate about the state of the economy, i.e. to \( \hat{\mu}_t \). Either they are able to perform the estimation on their own, or the estimation is done by a central authority in the economy, say in the form of the economic outlook by the central bank, which is then available to all agents. While the choice is not crucial, we favor the latter interpretation as being economically more meaningful.\(^{16} \) Finally, it is important to stress that while the economic outlook is an important piece of information, we assume that the actual consumption decision of agents will not exclusively depend on that estimate, but will be impacted by further factors, such as the local interaction.

**Social dynamics**

The important point where we deviate from the classic Lucas economy setup is the fact that we drop the representative agents assumption. In contrast, we consider an economy with a finite number \( N \) of agents that interact locally. In order to give meaning to the notion of *locality*, we need to impose some topology on our economy. From the variety of possible fixed networks, for the sake of

\(^{16} \) The choice can be justified with results in Carroll (2003), who shows that contrary to households, professional forecasters are able to form close to rational expectations.
simplicity, we choose the stable ring network. Thus, we can think of the $N$ agents to be situated along a line, where agent $N$ is connected to agent 1. More formally, let us define the neighborhood of agent $i$ as

$$N_i = \{j; (i-1) \mod N \leq j \leq (i+1) \mod N\}.$$ 

Note that $N_i$ is the set of the nearest adjacent agents on the ring, including agent $i$. At any point in time $t$, we let every agent assume one out of two different states, denoted as $S_i(t) = s$ with $s \in \{0, 1\}$. For our purposes, we assume that an agent $i$ in state $S_i = 1$ expects the economy to be in the high-growth regime, while in state $S_i = 0$ she expects low growth rates. Loosely speaking, the two possible states might be interpreted as agents being pessimistic or optimistic, bearish or bullish, in good or bad mood or some similar dichotomy.

Each agent can observe the state of the agents in his neighborhood, but not the states of agents in the whole economy. We denote the collective state of his neighborhood as $S_{n_i}(t)$. The state of his neighborhood at time $t$ will influence the state of the agent at the subsequent period $t + 1$. If we only allowed for local interaction in this form, we would be in a pure contagion model as e.g. analyzed in Morris (2000). However, we are interested in the interplay between local interaction and incomplete information about the regime of the economy. Therefore the state of a single agent will also be influenced by the estimate of the regime $\hat{\mu}$.

---

17 See Willitte (2006) for a review of network topologies. In general, results will depend on the choice of the specific network. Morris (2000) however has shown that there exist conditions for which contagion is possible in arbitrary network topologies. Since our focus is not on contagion results, the specific choice of networks is not of primary importance for our purposes.

18 On the one hand, it is uncommon to include the agent in his own neighborhood, since usually one assumes the condition of irreflexivity (i.e. no agent is a neighbor of his own), but on the other hand it has the advantage to avoid cyclical behavior which occurs if odd agents are in state 1 while even agents are in state 0.

19 An extension to finitely many different states would be straightforward.

20 Considering the impact of the mood of investors on asset returns is by no means an exotic issue, but has already been discussed in the behavioral finance literature. Saunders (1993) and Hirshleifer & Shumway (2003) for example find evidence that the weather has an significant influence on stock returns, the channel being the mood of investors.

21 Different choices are possible, but for simplicity, define the state of the neighborhood as the equally weighted average of the states of the agents in the neighborhood, i.e. $S_{n_i} = |N_i|^{-1} \sum_{k \in N_i} S_k$. 

---

8
as summarized in the economic outlook. Thus, at an abstract level, we can summarize the dynamic evolution of the local interaction as the following updating rule,

\[ S_i(t) = f(S_{N_i}(t-1), \hat{\mu}(t), \theta). \] (2)

The state of agent \( i \) at time \( t \) is a function of the state of his neighborhood at the previous time, and the current estimate about the regime. The parameter \( \theta \) is introduced as governing the weight given to the different influencing factors and will be crucial for our further analysis. Amongst others, \( \theta \) controls for the influence of the neighborhood on a single agent, i.e. the intensity of local interaction.\(^{23}\)

To make \( f \) in (2) operational, we assume the following functional form,

\[ S_i(t) = \begin{cases} 0 & \text{with probability } (1 - \pi_t) \\ 1 & \text{with probability } \pi_t \end{cases}, \]

where the probability \( \pi_t \) is determined as

\[ \pi_t = g(\theta_N S_{N_i}(t-1) + (1 - \theta_N) \hat{\mu}(t)). \] (3)

Note that the argument of \( g \) is a weighted sum between the state of the neighborhood and the global information, where \( 0 \leq \theta_N \leq 1 \) denotes the weight given to \( S_{N_i} \) and \( g \) is a mapping \( g : [0, 1] \rightarrow [0, 1] \), given by

\[ g(y) = \frac{1}{2} \operatorname{sgn}(2y - 1)|2y - 1|^\theta_g + \frac{1}{2}, \quad \text{for } 0 \leq y \leq 1, \] (4)

with \( \theta_g > 0 \). First, note that \( S_{N_i} \) as well as \( \hat{\mu} \) are bounded between 0 and 1. The function \( g \), which has some similarity to sigmoid functions, ensures that \( y \in [0, 1] \) is mapped to the unit interval, whereby the actual mapping depends on the parameter \( \theta_g \). Figure 1 illustrates function \( g \) for three different values for \( \theta_g \). If the information given by the regime-estimate and the state of the neighborhood indicates that there is a 70\% probability for being in the high

---

\(^{22}\) Note that we do not consider the influence of market prices. There are basically two reasons: On the one hand, it is far from clear what agents can infer from market prices in such an environment, and on the other hand, in the pure exchange-economy prices have no speculative role.

\(^{23}\) The interpretation is equivalent to the parameter \( c \) in Cont & Bouchaud (2000), which they define as “the willingness of agents to align their actions”.

---
state (i.e. \( \theta_N S_{N_i}(t - 1) + (1 - \theta_N) \hat{\mu}(t) = 0.7 \)), then for \( \theta_g < 1 \) (solid line) this
translates to a \( \pi \)-value of close to 1, for \( \theta_g = 1 \) (dashed line) it is identical to
0.7, and for \( \theta_g > 1 \) (dotted line), it is mapped to a \( \pi \)-value close to 0.5. For
the extreme cases we have for \( \theta_g \rightarrow 0, \pi \rightarrow 1 \{ \theta_N S_{N_i}(t - 1) + (1 - \theta_N) \hat{\mu}(t) > 1/2 \} \), while for
\( \theta_g \rightarrow \infty, \pi \rightarrow 1/2 \). Thus the parameter \( \theta_g \) controls for the extent to which

The agent relies in her decision on her external information incorporated in the
economic outlook (regime-estimate, \( \hat{\mu}(t) \)) and her neighborhood’s past beliefs
(\( S_{N_i(t - 1)} \)), or on her individual, idiosyncratic information. For high values of
\( \theta_g \), the agent’s state is dominated by idiosyncratic, internal randomness. Low
values of \( \theta_g \) indicate that the state is predominantly determined by external
influences. Note that there are two different external factors, whereby the state
of the neighborhood is a local effect, while the regime-estimate is a global effect.
Thus, we can succinctly summarize the controlling parameters in \( \theta = (\theta_g, \theta_N) \).

Note that there are three boundary parameter constellations possible. If \( \theta_g \rightarrow 0, \)
then each agent randomly chooses to be optimistic or pessimistic, and thus the
parameter \( \theta_N \) has no influence whatsoever. On the other hand, if \( \theta_g \rightarrow \infty, \)
then for \( \theta_N = 1 \), agents exclusively look at the beliefs of their neighbors and ignore
the information from the economic outlook. Vice versa, for \( \theta_N = 0 \), they ignore
the beliefs of their neighbors and exclusively rely on the global information about
the state of the economy. In Figure 2, the temporal evolution of agents’ beliefs
is illustrated for these three cases. On the vertical axis, the number of agents is
indicated, while the horizontal axis indicates time steps. Each agent corresponds

Figure 1: Graphical illustration of the function \( g(y) \)

The figure shows the function \( \pi = g(y) = 1/2 \text{sgn}(2y - 1) |2y - 1|^{\theta_g} + 1/2 \) for three different
choices of \( \theta_g \); \( \theta_g < 1 \) (solid line); \( \theta_g = 1 \) (dashed line); \( \theta_g > 1 \) (dotted line). As \( \theta_g \rightarrow \infty, \)
\( \pi \rightarrow 1/2 \), i.e. the state is determined independent of external influences, while for \( \theta_g \rightarrow 0, \)
\( \pi \rightarrow 1 \{ \theta_N S_{N_i(t - 1)} + (1 - \theta_N) \hat{\mu}(t) > 1/2 \} \), i.e. the state is exclusively determined by external influences.
to a cell, that can either be black or white, corresponding to being either in state 0 or 1. In the three panels of Figure 2, we consider \( n = 100 \) agents, which are vertically aligned in each of the columns. Let \( S(t) \) denote the vector of individual states at time \( t \), i.e. \( S(t) = (S_1(t), \ldots, S_n(t)) \). Then each column corresponds to \( S(t) \), i.e. represents the state of the agents at a given point in time. Moving from left to right, we observe \( S(0), S(1), \ldots, S(t) \), i.e. the states of the agents as they evolve through time. In short, the three panels are a graphical representation of the process \( S(\tau)_{0 \leq \tau \leq t} \).

**Figure 2: Dynamic evolution of agents’ states for boundary cases**

The three panels show the temporal evolution of agents’ states. The horizontal axis indicates time steps. For each time step, agents are aligned on the vertical axis. Each agent is either in state 0 or 1, represented as black or white cell. In the upper panel, \( \theta_g \rightarrow 0 \); i.e. each agent’s state is determined independently. In the middle panel, \( \theta_g \rightarrow \infty \) and \( \theta_N = 0 \), i.e. all agents rely exclusively on external information, and in the lower panel, \( \theta_g \rightarrow \infty \) and \( \theta_N = 1 \), i.e. all agents rely exclusively on the observation of their neighborhood. (Note that all three simulations are computed on the basis of the same vector of random numbers, i.e. by initializing the pseudo-random number generator at the same seed value.)

The upper panel displays \( S(\tau)_{0 \leq \tau \leq t} \) for \( \theta_g \rightarrow 0 \), i.e. all agents ignore external influences and their beliefs are determined by purely internal factors. The picture that emerges shows a homogenous fluctuation of white and black cells which may be characterized as noise. In the middle panel \( \theta_g \rightarrow \infty \) and \( \theta_N = 0 \), which means that agents rely exclusively on external information, and that they only look at the information in the economic outlook and ignore the beliefs of their neighbors.
The panel shows that for each time step, agents are identical in their states, i.e. $S_i(t) = S_j(t)$ for all $i,j$. As the economic outlook over time suggests that the state of the economy changes, agents collectively change their states. Thus, the picture that emerges is one of vertical stripes. Finally, in the lower panel, we chose $\theta_g \to \infty$ and $\theta_N = 1$, i.e. agents rely on external information, but they exclusively follow the beliefs of their neighbors. In that case, we observe robust clusters of pessimistic and optimistic agents, that do not change their beliefs over time. The picture is one of horizontal stripes.

While the extreme parameter cases serve to illustrate the driving forces of the model in isolation, it will be more interesting in the following to consider cases where these effects interact with each other. Figure 3 displays two examples.

**Figure 3: Dynamic evolution of agents’ states for intermediate cases**

The three panels show the temporal evolution of agents’ states. The horizontal axis indicates time steps. For each time step, agents are aligned on the vertical axis. Each agent is either in state 0 or 1, represented as black or white cell. In the upper panel, $\theta_g = 0.2$ and $\theta_N = 0.1$; while for the lower panel: $\theta_g = 0.5$ and $\theta_N = 0.8$. (Note that both simulations are computed on the basis of the same vector of random numbers, i.e. by initializing the pseudo-random number generator at the same seed value.)

In the upper panel, $\theta_g = 0.2$ and $\theta_N = 0.1$. With respect to their external information, agents put a much larger weight on the regime estimate than on the state of their neighborhood. Furthermore, external information is not the only influence, and agents’ decisions are also affected by idiosyncratic influences. In the lower panel, $\theta_g = 0.5$ and $\theta_N = 0.8$. In that case, agents put a much larger weight on the state of their neighborhood. The interaction of the different influences has important consequences. In particular in the lower panel, we observe

---

24 They are identical in each time step except for the initial random distribution of states in $t_0$. 

12
persistent clusters of optimistic/pessimistic agents, whose extent is subject to random fluctuations over time. Similar clusters are obtained in the upper panel, but due to the fact that the global information has a larger weight, these cluster fade out quickly, and a majority of agents switches states when new information with respect to the state of the economy is available.

Having specified the social dynamics in a heterogenous agents model, the next section discusses the market mechanism.

**Consumption decision, demand, and price formation**

The state of an agent is assumed to determine if she expects the economy to be in the high or low growth regime, which, on the aggregate level, will have an impact on the market clearing price. Conditional on her state, agents in our endowment economy try to maximize their life-time utility from consumption. We assume that agents in our economy are simple in the sense that they have preferences that can be ordered by a standard time-additive utility function over current consumption with constant relative risk aversion, i.e.

\[ u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}. \]  

Thus, with respect to life-time utility, every agent tries to maximize

\[ U(c) = \mathbb{E} \left( \sum_{j=0}^{\infty} \beta^j u(c_{t+j}) \right), \]

where \( \beta \) denotes the subjective time discount factor. In the sense of the Bellman equation, \( U(c) \) can be decomposed into current utility of consumption (as of time \( t \)) and some expected future utility, which is denoted as the value function \( V \), i.e.

\[ U(c) = u(c_t) + \beta V_{t+1}, \quad \text{where:} \quad V_{t+1} = \mathbb{E}_t \left( \sum_{j=1}^{\infty} \beta^{j-1} u(c_{t+j}) \right). \]  

We consider the case, where agents adopt a cautious or safety-first strategy in the following sense. For each \( t \), agents consider their value function to consist of the future expected utility which they obtain if they keep their portfolio unchanged. Let \( x_{i,t} \) denote the fraction of the risky asset that agent \( i \) holds at time \( t \), then
we take her value function \( V^{(s)}_{t+1} \) conditional on being in state \( S_i(t) = s \) to be

\[
V^{(s)}_{t+1} = \mathbb{E}_t \left( \sum_{j=1}^{\infty} \beta^{-1} u(x_t, D_{t+j}) | S_i(t) \right)
\]

\[
= \frac{x_t^{1-\gamma}}{1-\gamma} \sum_{j=1}^{\infty} \beta^{-1} \mathbb{E}_t \left( D_{t+j}^{1-\gamma} | S_i(t) \right).
\]

To call this a cautious or safety-first strategy is justified, because this is the (expected future) utility, agents can lock in today given their current information. Therefore, it represents a minimum or lower bound on the value function.\(^{25}\) Note that since we assume agents to lock in their future consumption possibilities, their decision does not depend on future price changes, which is consistent with our assumption that the state of an agent will not be influenced by market prices.

Taking into account the dividend dynamics (1) and the utility function (5), it is easy to show that

\[
V^{(s)}_{t+1} = \frac{(x_tD_t)^{1-\gamma} \tilde{\beta}(s)^{\cdot s} - 1}{1-\gamma}.
\]

where \( \tilde{\beta}(s) = \beta \exp \{ 1/2(1-\gamma)(2\mu(s) + \gamma \sigma^2_B) \} \}. \) Being in state \( s = 0 \), the agent expects the growth rate \( \mu(0) = \mu_i \), while for \( s = 1 \), \( \mu(1) = \mu_h \). By defining

\[
\gamma^{(s)} = \frac{\tilde{\beta}(s)^{\cdot s} - 1}{1-\beta(s)}
\]

this can also be written as

\[
V^{(s)}_{t+1} = u(x_tD_t) \cdot \gamma^{(s)},
\]

which shows that the value function equals current utility scaled by a factor \( \gamma^{(s)} \) which takes into account the expected dynamics of the dividend process and, more importantly, the anticipated state of the economy. If the economy is expected to be in the high state, the value function assumes a larger value than if the economy is in the low state, since \( \gamma^{(0)} < \gamma^{(1)} \).

The budget constraint for each agent requires that for any \( t \), consumption must equal the difference between the value of the portfolio before and after readjusting the asset holdings. Let \( x_t^- \) denote the fraction of the risky asset, i.e. the shares, right before readjustment, then the budget constraint for each agent is:

\[
c_t = (P_t + D_t) x_t^- - P_t x_t,
\]

\(^{25}\) This is also the reason why we ourselves were cautious in writing above, that agents try to maximize their life-time utility.
where obviously, $P_t$ is the market price of the risky asset. Plugging in the budget constraint and the value function in (6), the first-order condition for $\max_x U(c)$ is

$$u'(P_t + D_t x^-_t - P_t x_t) + \beta V(s) u'(x_t D_t) = 0,$$

Solving for $x_t$, we find

$$x_t = \frac{\tilde{V}(s) x^-_t (1 + \omega_t)}{\tilde{V}(s) \omega_t + \omega_t^{1/\gamma}}, \quad (7)$$

where $\omega_t = P_t / D_t$ is the price-dividend ratio, and $\tilde{V}(s) = (\beta V(s))^{1/\gamma}$. Note that we are able to explicitly derive a demand function for agents which depends on their state $S_i(t)$. Note that demand is a non-linear function of price and that demand depends on the state of the individual agent. Taking the first derivative of $x_t$ with respect to $\tilde{V}(s)$ yields

$$\frac{\partial x_t}{\partial \tilde{V}(s)} = \frac{x^-_t (1 + \omega_t) \omega_t^{1/\gamma}}{\left(\tilde{V}(s) \omega_t + \omega_t^{1/\gamma}\right)^2} > 0.$$

Together with the fact that

$$\frac{\partial V(s)}{\partial \mu} = \frac{(1 - \gamma) \tilde{b}(s)}{(\tilde{b}(s) - 1)^2},$$

it can be verified that for $\gamma > 1$;

$$x_t^{(1)} < x_t^{(0)}.$$

An agent being in state 1 has a lower demand for the risky asset. This follows from the well-established hedging demand in a Lucas economy. As long as $\gamma > 1$, i.e. hedging demand dominates the substitution effect, an agent who expects higher future growth rates, consumes more today and saves less. Vice versa, agents having pessimistic outlooks are saving more and consume less; thus their demand for the risky asset which is the only saving technology in this model increases.

To close the model, let $X_t^{(s)}$ be the aggregate demand of all agents in state $s$, i.e. $X_t^{(s)} = \sum_i x_i^{(s)}$, then the market clearing condition requires

$$1 = \sum_i X_t^{(s)}$$

15
for every $t$, from which the market clearing price can be determined numerically.

It is instructive to compare our model setup to established results from the representative agent framework. First, note that if we reduce our setup to the representative agent economy, then it must hold that $x_t = x_t^r = 1$, i.e. demand by the single representative agent must equal 1 in every period. Applying this restriction on (7), it follows that the price-dividend ratio simplifies to

$$\omega^{(s)} = \beta y^{(s)} = \frac{\beta^{(s)}}{1 - \beta^{(s)}}. \quad (8)$$

Conditional on a specific state (or growth rate), the price-dividend ratio is a constant and exactly the same as the one obtained by e.g. Mehra & Prescott (1985, 2003)\textsuperscript{26} and many other contributions within the representative agent paradigm. E.g. Cecchetti et al. (1990, 2000) derive the explicit formula for the price-dividend ratio under the assumption of a Markov regime-switching model. In their model, the price-dividend ratio is piece-wise constant conditional on the current state, and jumps to the alternative levels once the underlying Markov chain switches the state. As we show in a subsequent section, we obtain the same result if we assume that all agents decide exclusively on the basis of the regime-estimate $\hat{\mu}_t$.

3 Predictability

3.1 Empirical stylized facts for return predictability

Arguably, no other issue in financial economics has received more attention than the classic question: Are returns predictable? At first sight, evidence of predictability would seem like a serious blow against the efficient market hypothesis, and would suggest that there is a profitable trading opportunity out there. While early empirical papers which provided evidence for predictability were interpreted in that way,\textsuperscript{27} extensive subsequent work has shown that predictability is more likely to be the result of time variation in expected return, and thus not inconsistent with efficient markets. While this may reassure proponents of the efficient market hypothesis, it raises the obvious, but probably even more puzzling question: Where does time variation in expected returns come from? The literature

\textsuperscript{26} See e.g. equation (12) in Mehra & Prescott (2003).

\textsuperscript{27} As e.g. with Shiller (1984), Summers (1986).
has offered different explanations, including time-varying risk aversion, habit formation, long-run consumption risk or time-varying risk-sharing opportunities. However, there is still no conclusive evidence and consensus. Moreover, even the very existence of predictability is not a generally accepted fact. Doubts have been raised concerning the validity of the econometric tests being employed, and in particular Stambaugh (1999) has shown that regression coefficients can be heavily upward-biased and test statistics lean towards rejecting the null of no-predictability. Another line of attack claims that the evidence has only been established in-sample, while there is no out-of-sample evidence for predictability. Furthermore, it was argued that regression coefficients are unstable over time. The criticism spurred the use of more refined econometric approaches which in turn tend to confirm predictability. Lewellen (2004) for example shows that the bias correction developed in Stambaugh (1999) is too conservative in that it ignores useful information and he finds strong evidence for predictability once using this information. Lettau & Van Nieuwerburgh (2008) claim that it is important to take account of regime-shifts in the predictive variable and show that this mitigates coefficient instability. Dangl & Halling (2009) show within a Bayesian framework that allowing for dynamic coefficients provides strong evidence for predictability. Finally, Cochrane (2008) and Koijen & Van Binsbergen (2009) argue on the basis of a present value model and find evidence in favor of predictability. In particular, Cochrane (2008) argues that it is not sufficient to formulate a null of no return predictability, but the null must be a joint hypothesis due to the present value accounting relation. Evaluating the joint null hypothesis, he finds stronger evidence against the null.

As explanatory variable, one of the most frequently used predictors is the price-dividend ratio. Other financial ratios have been tested, as for example the price-earnings, book-to-market, consumption-to-wealth ratio or the short rate. Within our model setup it is most natural to focus on the price-dividend ratio as predictive variable. A corresponding regression equation can be motivated by the following reasoning. Start with the definition of time \( t \) gross stock return \( R_t \equiv (P_t + D_t)/P_{t-1} \) and the obvious identity

\[
1 = R_{t+1}^{-1} R_{t+1} = R_{t+1}^{-1} \frac{P_{t+1} + D_{t+1}}{P_t}.
\]

\[28\] For an overview, see e.g. Cochrane (2005), or Koijen & Van Nieuwerburgh (2009).

\[29\] See e.g. Goyal & Welch (2003, 2008).

\[30\] See e.g. Ang & Bekaert (2007)
Multiply by \( P_t/D_t \) and rearrange to get
\[
\frac{P_t}{D_t} = R_{t+1}^{-1} \left( 1 + \frac{P_{t+1}}{D_{t+1}} \right) \frac{D_{t+1}}{D_t}.
\]

Taking logs, using lowercase letters for log variables, and define \( pd_t \) as log price-dividend ratio \( pd_t = \ln(P_t/D_t) = p_t - d_t \) yields
\[
pd_t = \Delta d_{t+1} - r_{t+1} + \ln(1 + e^{pd_{t+1}})
\]
which is non-linear. Campbell & Shiller (1989) propose a linearization by taking a Taylor expansion about the point \( pd = \ln(P/D) \) (where \( P \) and \( D \) are the long-run means) to obtain
\[
pd_t = \Delta d_{t+1} - r_{t+1} + \rho pd_{t+1} + \kappa,
\]
where
\[
\rho = \frac{e^{pd}}{1 + e^{pd}}, \quad \kappa = \ln(1 + e^{pd}) - \rho pd
\]
Iterating forward, together with the transversality (or ‘no bubble’) condition \( \lim_{t \to \infty} \rho^t pd_t = 0 \), yields
\[
pd_t = \sum_{s=1}^{\infty} \rho^{s-1} \mathbb{E} (\Delta d_{t+s} - r_{t+s}) + \text{const.}
\]
(10)
Adding the expectation operator is justified, since the equation holds both ex post and ex ante. Equation (10) is central to the return predictability literature since it directly shows the link between future dividend growth, returns and the price-dividend ratio. From (10), a high \( \rho t \) must be followed by either high future dividend growth, low future returns or both. Thus, the predictability literature has largely focused on testing the following related three regression equations
\[
\begin{align*}
r_{t+1} &= \alpha_r + \beta_r dp_t + \epsilon_{t+1}^r, \\
\Delta d_{t+1} &= \alpha_d + \beta_d dp_t + \epsilon_{t+1}^d, \\
dp_{t+1} &= \alpha_{dp} + \phi dp_t + \epsilon_{t+1}^{dp}.
\end{align*}
\]
Although being still subject to ongoing discussion, the following stylized facts seem to be supported by the predictability literature: (i) Future returns seem to be predictable, (ii) Coefficient estimates increase with forecast horizons, (iii)

\[ 31 \text{ It is common to regress on the log dividend-price ratio } dp_t = -pd_t. \]
Absence of dividend growth predictability is additional evidence for return predictability, and (iv) Dividend-price ratios are highly persistent.

(i) There is evidence for return predictability, although coefficients are instable and the statistical significance is not overwhelming. Table 1 documents some selected results.

Most studies find a regression coefficient on the order of 0.1, and a $t$-statistic roughly around 2. Exceptions are Stambaugh (1999) and Lewellen (2004). It has to be noted, that reported regression results are derived for different samples across studies, and depend upon the length of the sample. Many studies report different coefficient estimates for various subsamples.

(ii) Regressions are usually done for data on a quarterly or yearly basis. As the time horizon is increased, evidence for predictability seems to be stronger. This is illustrated by selected results shown in Table 2. As the time horizon increases, the regression coefficients rise substantially accompanied by a sensible increase in $R^2$-numbers. However, although larger coefficient estimates and $R^2$ indicate stronger statistical evidence, this is not necessarily the case. It has been recognized that the increasing long-term coefficients can result from small short-term coefficients if the price-dividend ratio displays substantial persistence (like it does). In that case, predictability ’sums up’ over time. Boudoukh et al.

Table 1: Empirical results from return forecast regressions.

The table shows selected results from the empirical predictability literature for the return predictability regression. The first column reports regression coefficients, the second and third column indicate $t$-statistics and $R^2$'s.

<table>
<thead>
<tr>
<th>Return predictability regressions $r_{t+1} = \alpha_t + \beta_t \cdot dp_t + \epsilon_{t+1}$</th>
<th>$\beta_t$</th>
<th>$t$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.097</td>
<td>1.92</td>
<td>0.04</td>
<td>Cochrane (2008)</td>
</tr>
<tr>
<td>0.1</td>
<td>1.76</td>
<td>n.r.</td>
<td>Ang &amp; Liu (2007)</td>
</tr>
<tr>
<td>0.09</td>
<td>2.04</td>
<td>&lt;0.16</td>
<td>Lettau &amp; Van Nieuwerburgh (2008)</td>
</tr>
<tr>
<td>0.08 - 0.11</td>
<td>~2</td>
<td>n.r.</td>
<td>Ang &amp; Bekaert (2007)</td>
</tr>
<tr>
<td>0.14</td>
<td>0.17*</td>
<td>n.r.</td>
<td>Stambaugh (1999) (* p-value for $\beta = 0$)</td>
</tr>
<tr>
<td>0.66</td>
<td>4.6</td>
<td>0.004</td>
<td>Lewellen (2004)</td>
</tr>
</tbody>
</table>
Table 2: Empirical results for long-horizon regressions.

The table shows selected results for long-horizon return predictability regressions. Column 2-5 report regression coefficients and $R^2$s for a time horizon of 1 to 5 years. Results suggest that the evidence for predictability strengthens as the horizon is increased.

<table>
<thead>
<tr>
<th>Long-horizon return predictability regressions, $\tau$ years ahead. $r_{t+1} + \ldots + r_{t+\tau} = \alpha_r + \beta_r dp_t + \epsilon_{t+\tau}$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_r$</td>
<td>0.19</td>
<td>0.38</td>
<td>0.53</td>
<td>0.65</td>
<td>n.r.</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.07</td>
<td>0.14</td>
<td>0.21</td>
<td>0.27</td>
<td>n.r.</td>
</tr>
<tr>
<td>$\beta_r$</td>
<td>0.13</td>
<td>0.24</td>
<td>0.27</td>
<td>0.3</td>
<td>0.76$^*$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.04</td>
<td>0.06</td>
<td>0.07</td>
<td>0.07</td>
<td>0.25$^<em>$$^</em>$ for $\tau=6$</td>
</tr>
<tr>
<td>$\beta_r$</td>
<td>0.13</td>
<td>0.25</td>
<td>0.39</td>
<td>0.46</td>
<td>0.52</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.05</td>
<td>0.1</td>
<td>0.16</td>
<td>0.18</td>
<td>0.21</td>
</tr>
</tbody>
</table>

(2008) elaborate on this point and show that long-horizon results present hardly any additional independent evidence about predictability. Still, although there may be no additional statistical significance, results of long-horizon predictability underscore the important economic significance that long-horizon forecasts can predict roughly one third of the variation in future returns.

(iii) From (10), dividend-price ratios predict either returns, dividend growth or both. Regressing $dp_t$ on dividend growth as in regression equation (12), it has been found that the null of $\beta_d = 0$ is hard to reject. There seems to be no evidence for predictability of future dividend growth. Fewer papers report results on regressions as in (12). Lettau & Van Nieuwerburgh (2008) report a $\beta_d$ coefficient between 0.005 and 0.035 and $t$-statistics smaller than 0.33. Similarly, Cochrane (2008) finds $\beta_d = 0.008$ with $t = 0.18$, which is way beyond from being significant.

Further evidence comes from the variance decomposition of the price-dividend ratio. Note that from (10) it follows that

$$\text{var}(dp_t) = \text{cov}(dp_t, \Delta d_\infty) - \text{cov}(dp_t, r_\infty),$$

where $\Delta d_\infty = \sum_{s=1}^{\infty} \rho^{s-1} \Delta d_{t+s}$ and $r_\infty = \sum_{s=1}^{\infty} \rho^{s-1} r_{t+s}$. Cochrane (1992) reports estimates

$$\frac{\text{cov}(dp_t, \Delta d_\tau)}{\text{var}(dp_t)} = -0.34,$$

$$\frac{\text{cov}(dp_t, r_\tau)}{\text{var}(dp_t)} = 1.37.$$
for $T = 15$. Thus, showing that the variance in price-dividend ratios is largely influenced by variations in expected returns rather than revisions of expected dividend growth.

(iv) Dividend-price ratios display significant persistence. From (13), $dp_t$ follows an AR(1) process. Estimates for the autoregression coefficient $\phi$ tend to be large, and close to a unit root. The data seems not to be generated by a stable process. Lewellen (2004) reports autoregression coefficients as large as $\phi = 0.999$ for log dividend yields in the earlier part of his sample (1973-2000) and $\phi = 0.997$ for the entire sample. Lettau & Van Nieuwerburgh (2008) estimate $\phi = 0.945$ for the entire sample under the assumption that there is no structural break. In Cochrane (2008), he finds $\phi = 0.941$, whereas in Cochrane (2005) he reports values for $\phi$ in the range 0.82 to 0.97, depending on the time period of the sample. Finally, Goyal & Welch (2003) report $\phi = 0.847$ for a yearly lag, and show that for a 5 and 10 year forecast horizon this drops down to 0.47 and 0.25 respectively.

3.2 Evidence on predictability from the simulated economy

Obviously, the most basic representative agent models are not able to explain the empirical evidence. As shown in (8), the price-dividend ratio in an economy without regime-shifts is a constant. Taking into account regime-shifts in the growth rate of the endowment process as in Cecchetti et al. (1990, 2000), makes the price-dividend ratio piecewise constant conditional on the state. This temporal pattern is far from what can be observed.

In an attempt to explain the predictability evidence, the theoretical asset pricing literature can be classified into two broad approaches. On the one hand, models have been put forward that retain the assumption of a simple (i.e. constant coefficients, persistent unit root) endowment process, but consider a more complex representative agent. Additional complexity can be achieved by assuming a more sophisticated utility function which displays time-varying preferences, time-varying risk aversion or habit-formation. Standard power utility preferences are frequently replaced by Epstein-Zin/Weil preferences, which allow for a separation of the intertemporal substitution and risk aversion effect. Combining a simple

\[32\text{ See e.g. Menzly et al. (2004) for time-varying preference, Campbell & Cochrane (1999) or Collard et al. (2006) for habit formation.}\]
endowment process with a complex representative agent may generate the complex time series properties that lead to predictability. On the other hand, the second approach is to consider a standard simple representative agent, but to assume that the endowment process is more complex. For example, Calvet & Fisher (2007) consider a dividend process with state-dependent volatility, and assume that the state is controlled by a Markov-switching multifractal process with different frequencies. Bansal & Yaron (2004) specify a dividend process whose growth rate follows an AR(1) process which introduces a persistent predictable component. This way, an already complex endowment process feeds into a simple representative agent, which in turn yields a complex price-dividend time series.

Our model offers a third, alternative way for explaining predictability. We assume both, a simple endowment process as well as simple agents. However, we do not have a single representative agent, but consider a large number of (otherwise identical) agents that interact locally. Our results show that allowing for social dynamics in the sense of neighborhood effects is surprisingly successful in explaining the empirical evidence.

To fix ideas, Figure 4 shows the simulated log dividend-price ratio \((dp_t)\) from the model outlined in section 2 for three different parameter sets. For \(\theta_g = 0\) (dotted line), all agents behave the same way. \(dp_t\) is piecewise constant and jumps as the economy is supposed to switch to the alternative regime. For \(\theta_N = 1, \theta_g = 0.4\) (lower dashed line), agents put a high weight on the beliefs of their neighbors. The resulting dividend-price ratio displays persistent behavior. On the contrary, for \(\theta_N = 0.4, \theta_g = 1\) (middle solid line), the \(dp_t\) series seems to follow a stationary process. The visual impression can be confirmed by a formal test. The solid line has an autocorrelation coefficient of 0.87 and the null of a unit root can be strongly rejected. In contrast, the lower dashed \(dp_t\) series (with a high degree of local interaction) has an autocorrelation of 0.994 and the of a unit root cannot be rejected.

The subsequent section analyzes in more detail if local interaction can account for the predictability evidence, and if so under which conditions.

**Short-horizon return predictability**

For the simulation of our economy, we use the following parameter set, unless stated otherwise. \(\mu_h = 0.0225, \mu_l = 0.015, \sigma_D = 0.05\) on an annual basis. The
Figure 4: Simulated log dividend-price ratio series.

The figure shows realizations of the log dividend yield series for three parameter constellations. For $\theta_N = 0$ (dotted line), $dp_t$ is piecewise constant. For $\theta_N = 1, \theta_G = 0.4$ (lower dashed line), $dp_t$ displays large persistence. For $\theta_N = 0.4, \theta_G = 1$ (middle solid line), $dp_t$ seems stationary. The table on the right side confirms the visual inspection by an Augmented Dickey-Fuller test for a unit root. (Note that the three $dp_t$ realizations are computed on the basis of the same vector of random numbers, i.e. by initializing the pseudo-random number generator at the same seed value.)

<table>
<thead>
<tr>
<th>$\theta_N$</th>
<th>0.4</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_G$</td>
<td>1</td>
<td>0.4</td>
</tr>
</tbody>
</table>

$$dp_{t+1} = c + \phi dp_t + \gamma dp_{t+1}$$

| $\phi$ | 0.867 | 0.904 |
| $t$    | 38.6  | 410.6 |

Augmented Dickey-Fuller

$dp_t$ has a unit root

$p$-value 0.017 0.221

The dividend process is initially normalized to log $D_0 = 1$. Time steps are assumed to be months, i.e. $\Delta t=1/12$. We simulate $T = 600$ time steps, which corresponds to 50 years of data. The transition probability is symmetric, i.e. $p_{h,t} = p_{t,h} = 0.02$. Therefore, the growth rate is approximately expected to change once every 4 years, which roughly corresponds to the length of a business cycle.

The economy consists of $N = 500$ agents. For each agent, the subjective discount factor is $\beta = 0.995$, i.e. 0.95 per year and the risk aversion parameter is set to $\gamma = 3$. Note that although our focus is on qualitative results, this calibration is consistent with empirical evidence as reported in e.g. Mehra & Prescott (2003). With respect to the choice of $\theta_N$ and $\theta_G$, we cannot rely on empirical evidence. In this section, we present results for $\theta_N = 0.8$ and $\theta_G = 0.5$. We show in section 3.3 that this calibration of the model is a reasonable choice. Panels on the left side of Figure 5 show regression results for one simulation run, and report corresponding statistics. On the right side, we plot regression coefficients on the horizontal axis.

33 The transition probabilities correspond roughly to estimates in Cecchetti et al. (1990, 2000) and Ozoguz (2009). However, they find that there is a substantial asymmetry, in that the probability to stay in an expansion is much higher than to stay in a recession.

34 See also Cochrane (2005), p. 455f for a discussion.
Figure 5: Regression results from the simulated economy.

In the left column, regression results for a single simulation run are reported, while the right column reports average results over \( K = 10,000 \) simulation runs. The first two rows contain results from predictive regressions on expected returns and dividend growth. The lower two panels report results for autocorrelation in the dividend yield and return series. Regressions are performed with de-meaned series. Reported standard errors are Newey-West corrected for heteroskedasticity. The right panels report standard deviation from the simulated distribution of the regression coefficient \((\sigma(\cdot))\), and average (asymptotic) \( t \)-statistics \((\bar{t})\). The parameter choice is \( \theta_\alpha = 0.8 \) and \( \theta_\beta = 0.5 \).
against \( t \)-statistics on the vertical axis for \( K = 10,000 \) simulation runs.\(^{35}\)

In the first, upper row, results are shown for the basic predictive regression from (11). The plot on the left side shows that the relationship is not very strong but positive. The estimated regression coefficient is \( \hat{\beta}_r = 0.036 \) with a Newey-West corrected standard error of 0.0179 giving a \( t \)-statistic of 2.018, which yields a \( p \)-value of 4.4\%.\(^{36}\) An \( R^2 \) of 0.65\% confirms the modest explanatory power and implies that the current dividend-price ratio is able to predict less than 1\% of the variation in next months returns. These results are drawn from a single path of the simulated economy. In the right panel we report regression coefficients for \( K = 10,000 \) simulation runs. The average regression coefficient is somewhat smaller at 0.0342, with a standard deviation of 0.0138. Interestingly, note that the \( t \)-value from the simulated distribution is 2.47, which is larger than the average asymptotic \( t \)-value of 2.192. Therefore, there is evidence for predictability albeit at a statistically modest level.

In the second row, similar results are reported for the dividend growth regression as in (12). Regression coefficients are not significantly different from zero. Thus, there is no evidence that current dividend-price ratios can predict future changes in the dividend growth rate. The third row estimates the autocorrelation in log dividend yields. From the single simulation run, we find an autocorrelation coefficient of 0.983 with an associated \( t \)-value which is beyond any statistical doubt. Averaging over repeated simulations, the autocorrelation is slightly higher at 0.987.

Finally, the last, lower row reports results from regressing returns on one-month lagged returns. The left panel reports the coefficient from the regression

\[
\hat{r}_{t+1} = \alpha_{ac} + \psi \hat{r}_t + \epsilon_{t+1}^{ac}.
\]

The estimated coefficient \( \hat{\psi} \) is slightly negative but the standard deviation of the simulated coefficient distribution is roughly four times as large, thus being far from statistically significant. Thus, current returns cannot predict future returns, which is well in line with empirical evidence, and consistent with the usual idea of efficient markets.

\(^{35}\) For visibility, only 2,000 data points are plotted.

\(^{36}\) Note that the correction for heteroskedasticity does not affect the results in a substantial way. The unadjusted standard error is 0.0183, with a \( t \)-stat of 1.98, yielding a \( p \)-value of 4.83\%.
Long-horizon return predictability

Results for the one-period, i.e., one-month ahead predictive regression suggest significant, but modest predictability. As discussed in the previous section, regression coefficients, $t$-stats and $R^2$ tend to increase as the time horizon is enlarged. We perform regressions for our simulated economy by regressing $\tau$-month ahead returns $r_{t,\tau} = r_{t+1} + \ldots + r_{t+\tau}$ on the current dividend yield. Table 3 reports results. As the forecast horizon is enlarged, the (average) regression coefficient increases substantially. On a quarterly basis, the coefficient is 0.127 which is well in line with previous results. As we consider longer horizons, the coefficient increases fast, reaching levels of around 1 for three years ahead returns. The statistical significance seems to increase as well, as documented by the asymptotic $t$-statistics, which are slightly above 2 for short horizons, and increase substantially as the time horizon is enlarged, reaching double-digit levels. However, the increasing statistical significance has to be interpreted with care, as Boudoukh et al. (2008) have shown that long-horizon regressions do not necessarily provide further independent evidence from short-horizon regression. Actually, this can be confirmed from our results, when we compute the standard deviation of the forecast coefficients ($\sigma(\hat{\beta}_r)$) from $K = 1,000$ simulation runs. The empirical standard deviation increases proportionally, yielding a $p$-value around 1% for

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_r$</td>
<td>0.0341</td>
<td>0.0654</td>
<td>0.1273</td>
<td>0.1893</td>
<td>0.3753</td>
<td>0.7206</td>
<td>1.008</td>
<td>1.3766</td>
</tr>
<tr>
<td>$t$ (asymp.)</td>
<td>2.18</td>
<td>3.04</td>
<td>4.31</td>
<td>5.34</td>
<td>7.85</td>
<td>11.22</td>
<td>12.83</td>
<td>12.36</td>
</tr>
<tr>
<td>$\sigma(\hat{\beta}_r)$</td>
<td>0.0138</td>
<td>0.0282</td>
<td>0.0567</td>
<td>0.0846</td>
<td>0.1653</td>
<td>0.3243</td>
<td>0.4832</td>
<td>0.8212</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.0071</td>
<td>0.0103</td>
<td>0.0125</td>
<td>0.0128</td>
<td>0.0118</td>
<td>0.0133</td>
<td>0.0186</td>
<td>0.0471</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0101</td>
<td>0.0195</td>
<td>0.0383</td>
<td>0.0571</td>
<td>0.1105</td>
<td>0.1929</td>
<td>0.2356</td>
<td>0.2357</td>
</tr>
</tbody>
</table>

Table 3: Regression results for long-horizon predictability.
nearly all time horizons. However, even when the statistical evidence does not strengthen the case for long-horizon predictability, the results demonstrate the strong economic evidence for long-horizon prediction as documented by the substantial increase in the fraction of explained variation. While for the one-month ahead prediction only roughly 1% of the variation in future returns can be explained, this number increases to around 20% for long-horizons.

Instability of regression coefficients

The evidence for return predictability has been challenged by the finding that the forecasting relationship is not very stable over time. Various empirical studies have subdivided their sample in different subsamples and found that the estimated regression coefficients vary substantially across time periods. In particular, Paye & Timmermann (2006) and Lettau & Van Nieuwerburgh (2008) address the issue in more detail and reject the hypothesis of a constant regression coefficient. Lettau & Van Nieuwerburgh (2008) estimate return coefficients for a 30-year rolling window and document that the coefficient fluctuates between 0 and 0.5. In line with their approach, we estimate predictive regressions for a 25-year rolling window for our simulated economy. Figure 6 reports results.

Figure 6: Time variation in estimated regression coefficients.

The left panel shows the regression coefficient estimated from a 300 time steps rolling window, rolled over one 600 time step simulation run. The right panel shows the average (solid line) rolling window estimate over \( K = 10,000 \) repetitions, together with one standard deviation bounds (dashed lines). The parameter choice is \( \theta_N = 0.8 \) and \( \theta_g = 0.5 \).

The left panel shows \( \hat{\beta}_r \) estimated from a 25-year (300 time steps) window, which is rolled over the 50 year (600 time steps) sample from a single simulation
run. In the right panel we repeat $K = 10,000$ simulation runs and report the average rolling window estimates, together with one standard deviation bounds (dashed lines). In particular from the left panel we observe substantial time variation in the forecasting relationship. For the early part of the sample, the prediction coefficient is around 0.06, then drops down to zero and recovers to levels around 0.08. The substantial variation is confirmed by the right panel, which displays one standard deviation bounds around the average regression coefficient. Thus, the model confirms coefficient instability over time and supports the approach to estimate time-varying coefficient regression models such as e.g. in Dangl & Halling (2009). The significant time variation in the forecasting relationship has been interpreted as an explanation why in general out-of-sample prediction has performed only poorly.\textsuperscript{37} From unreported results, we confirm that out-of-sample tests perform poorly in our simulations. Our findings strengthen the conclusion made by Lettau & Van Nieuwerburgh (2008) that the crucial part in out-of-sample prediction is the detection of shifts in the underlying forecasting relationship.

**Evidence from a joint hypothesis test**

The evidence for return predictability so far is there, but is not overwhelming. Short-horizon regressions deliver significant results, although $t$-stats are modest and there is evidence for high coefficient instability. However, proponents of return predictability have developed more powerful statistical tests in an attempt to strengthen the evidence. One way is to go back to equation (10), which states that variation in the dividend-price ratio must be due to changes in returns or dividend growth or both. Since the dividend-price ratio can be observed to vary over time, it necessarily follows that there need to be either variation in expected returns, expected dividend growth or both. In other words: If returns are assumed to be not predictable, then dividend growth must be predictable. Cochrane (2008) uses this implication to construct a joint hypothesis test from which he concludes that actually the absence of dividend growth predictability gives even stronger evidence for return predictability. We follow his approach to test for predictability in our simulated economy.

To understand the argument, note that from the Vector Autoregressive (VAR) representation in equations (11)-(13), and the log-linearization of returns in (9) it follows that the regression coefficients are related by the following approximate identity

$$\beta_d = \rho \phi - 1 + \beta_r. \quad (14)$$

Similarly, regression errors are related by

$$\epsilon_{t+1}^r = \epsilon_{t+1}^d - \rho \epsilon_{t+1}^{dp}.$$ 

From the log-linearization in (9), $\rho$ is defined as $\rho = \frac{1}{\epsilon^d_{t+1} + \epsilon^{dp}_{t+1}}$, where $\bar{pd}$ is the mean log dividend-price ratio. In our simulated sample, we find $\bar{pd} = 3.367$ or $\rho = 0.9667$, and an autocorrelation coefficient of $\phi = 0.9870$. Under the null of no return predictability, $\beta_r = 0$, and the restriction from the present value model in (14) implies that $\beta_d$ cannot be null, and has to be negative, i.e. $\beta_d = -0.0459$. Therefore, the hypothesis of no return predictability cannot be tested by a null of only $\beta_r = 0$, but the null needs to be a joint hypothesis of $\beta_r = 0$ together with $\beta_d < 0$. Thus, under the restrictions from the present value model, the VAR under the null must take the form

$$
\begin{pmatrix}
    dp_{t+1} \\
    \Delta d_{t+1} \\
    r_{t+1}
\end{pmatrix}
= 
\begin{pmatrix}
    \phi \\
    \rho \phi - 1 \\
    0
\end{pmatrix}
\cdot dp_t + 
\begin{pmatrix}
    \epsilon_{t+1}^{dp} \\
    \epsilon_{t+1}^d \\
    \epsilon_{t+1}^d - \rho \epsilon_{t+1}^{dp}
\end{pmatrix}. \quad (15)
$$

To evaluate the null, the system in (15) is simulated by using the sample estimates from our simulated economy. Error terms $\epsilon_{t}^{dp}$ and $\epsilon_{t}^d$ are taken to be bivariate normal with a covariance matrix estimated from the simulated sample. We simulate (15) over $T = 600$ time steps and estimate the regression coefficients with that sample. We repeat the simulation $K = 10,000$ times. Figure 7 reports results, where for visibility reasons, only 2000 data points are plotted.

In the left panel, we plot the regression coefficients $\beta_r$ on the horizontal axis, and $\beta_d$ on the vertical axis. The solid lines are drawn at the sample estimates from section 3.2, i.e. $\hat{\beta}_r = 0.0342$ and $\hat{\beta}_d = -0.012$, while the dashed line indicate the coefficients implied by the null. Therefore, the white point inside the point

38 $\bar{pd} = 3.367$ implies a price-dividend ratio of 28.9 or a dividend yield of 3.45%.

Figure 7: Simulated joint distribution of regression coefficients.

The figure shows the simulated joint distribution of return ($\beta_r$) and dividend-growth ($\beta_d$) predictability coefficients (left panel) and $t$-statistics (right panel) from the null hypothesis in (15). Solid lines indicate the sample estimates from section 3.2; dashed lines indicate the coefficient values implied by the null. Percentage values indicate the frequency of simulated realizations larger than the sample estimate for $K = 10,000$ simulation runs. (For visibility reasons the figures are plotted with only 2,000 data points.)

Cloud represents the null, and is drawn at $\beta_r = 0$ and $\beta_d = -0.0459$ which follows from the restriction discussed above. From the left panel we observe that the univariate evidence with respect to return predictability is weak. There are 10.4% of the simulation runs which are greater than the sample estimate, from which we would not be able to reject the null of $\beta_r = 0$ with much confidence. However, the evidence with respect to (the absence of) dividend growth predictability is stronger. Only 3.9% of the simulation runs have values larger than the sample $\beta_d = -0.012$. These results indicate, that as in Cochrane (2008), the absence of dividend growth predictability is a stronger statistical piece of evidence for return predictability, than the finding of a positive $\beta_r$ alone.\textsuperscript{40} The right panel plots asymptotic $t$-values. A similar pattern can be observed, but there is no big difference between $t_r$ and $t_d$ in the frequency with which simulation runs lie outside the sample $t$-statistics. Taken together, we find that within our simulated economy, similar patterns of predictability can be observed as in real data, where the absence of dividend growth predictability is strengthening the evidence that returns are predictable.

\textsuperscript{40} Cochrane (2008) calls this finding the ‘dog that did not bark’ effect, following the famous Sherlock Holme’s case.
3.3 The role of social dynamics

Results in the previous sections were derived for a parameter constellation of $\theta_N = 0.8$ and $\theta_g = 0.5$, i.e. a high degree of local interaction and a modest degree of idiosyncratic influences. We were arguing that this parameter choice is reasonable. In this section, we discuss and justify the parameter choice. To this end, we average over repeated simulation runs of our economy for various parameter constellation ($\theta_N, \theta_g$) and compute the four crucial regression coefficients: $\beta_r$, $\beta_d$, $\phi$ and $\psi$, i.e. return, and dividend growth prediction coefficients, and the autocorrelation in the dividend-price ratio as well as in returns. Figure 8 displays the corresponding results for $K$ repetitions. Plots on the left side report the regression coefficients, while the panels on the right side show $t$-statistics. In each panel, $\theta_N$ is on the horizontal axis, and three graphs are drawn for $\theta_g = 0.5$ (solid line), $\theta_g = 1$ (dashed line), and $\theta_g = 1.5$ (dotted line).

Results for dividend growth are omitted, since we do not find significant estimates in any of the parameter constellations. To understand the results in Figure 8, start by observing the behavior of the regression coefficients for the case $\theta_g = 1.5$, i.e. the dotted lines. For a broad range of $\theta_N$-values, the return coefficient is extremely large and highly significant. However, by looking at the autocorrelation in the dividend-price ratio and returns, we observe that the high return forecast coefficient comes along with a very small $\phi$, i.e. no or only weak autorecorrelation in the dividend yield on the one hand, and large (and highly significant) negative autocorrelations in returns on the other hand. Observing these three facts simultaneously contradicts empirical evidence. Neither are dividend yields close to stationary, nor do we observe large negative autocorrelations in returns. Thus, this does not seem to be a reasonable calibration for our simulated economy. Now, as $\theta_g$ is decreased (i.e. agents rely more on external influences), the extent of the return predictability coefficient $\beta_r$ diminishes substantially, approaching a reasonable level of magnitude. For the dashed line ($\theta_g = 1$), $\beta_r$ assumes values around 0.2, and decreases as agents interact more strongly. The smaller return coefficient comes along with a substantially higher autocorrelation in the dividend yield. Furthermore, $\phi$ increases in $\theta_N$, which is in line with economic intuition. As agents put a large weight on the beliefs of their neighborhood in the past time step, this introduces persistence in the evolution of the aggregate beliefs in the economy which in turn results in persistence in the price-dividend ratio. With respect to the autocorrelation in returns, $\psi$ assumes large (highly
Figure 8: Regression results for different levels of social dynamics

Panels on the left side report regression estimates for return predictability and autocorrelations in dividend yields and returns against variation in the weight on social dynamics ($\theta_N$). The right panels report corresponding asymptotic Newey-West corrected t-values. In each panel three graphs are plotted, corresponding to $\theta_g = 0.5$ (solid line), $\theta_g = 1$ (dashed line), and $\theta_g = 1.5$ (dotted line).

significant) negative values for $\theta_N$ being low, but displays a steady increase, to the extent of becoming insignificant for $\theta_N$ close to one. A similar, albeit more pronounced, observation can be made for the case plotted as solid lines, i.e. for $\theta_g = 0.5$. Return predictability coefficients are smaller in magnitude but still significant, although in some cases t-statistics are hardly above 2. Persistence in dividend yields is even higher, approaching near unit-root, and return autocorrelations are slightly negative but most often insignificant.
Therefore, the central insight to be gained from the analysis can be summarized as follows. As long as agents’ consumption decisions are predominantly influenced by idiosyncratic influences, the price-dividend ratio follows a stationary distribution. Stationarity, in turn, induces significant negative autocorrelations in returns on first lags,\(^{41}\) which are responsible for predictability in returns from the current price-dividend ratio.\(^ {42}\) As agents rely more heavily on external influences, the price-dividend ratio displays more persistence. For the case that agents put a large weight on the global information about the state of the economy, the price-dividend ratio tends to jump between two levels. For the case that agents put a large weight on the beliefs of their neighborhood, slow-moving changes in aggregate beliefs with a high degree of persistence are obtained. Thus, within our artificial economy, a slow-moving, persistent price-dividend ratio is only consistent with a high degree of local interaction.

Taken together, the numerical analysis suggests that on the one hand high persistence levels in dividend-price ratios are only possible for low values of \( \theta_g \), i.e. for cases where agents are largely influenced by external factors. On the other hand, the absence of significant return autocorrelations as well as reasonable \( t \)-statistics for \( \beta_r \) can only be found for high \( \theta_{\beta} \)-values, i.e. for cases where the external influences are predominately given by neighborhood effects. These findings explain and justify the parameter choice made in the previous sections. Only the combination of a large weight on external influences (i.e. small values for \( \theta_g \)) together with the fact that these external influences are mainly due to local interaction effects (i.e. large values for \( \theta_{\beta} \)) yields simultaneously numerous results that are consistent with empirical evidence. As it turns out, an appropriately calibrated model is surprisingly successful in matching many of the stylized facts which have been established in the return predictability literature.

4 Conclusion

Predictability of asset returns is a long-standing and controversial topic. Although being subject to ongoing discussions, the empirical evidence seems to suggest that there is some predictability which is due to time-varying expected returns, thus being not at odds with the notion of efficient markets. Within

\(^{41}\) This finding is consistent with the results in Cecchetti et al. (1990).

\(^{42}\) Indeed, from unreported results, we find that the partial autocorrelation function at higher lags goes down to zero.
the theoretical asset pricing literature, attempts to explain predictability can be
classified into two broad strands. The first approach consists in retaining the
assumption of a simple underlying endowment process, but to assume that the
representative agent behaves in a more complex way (for example due to habit
formation, time-varying risk preferences, etc.). A second approach is to retain
the assumption of a simple representative agent, but to consider a more complex
endowment process (for example by incorporating stochastic volatility, or by link-
ing the endowment process to other state variables). Both ways, the resulting
time series for dividend yields and returns may display the observed empirical
patterns of predictability.

In this article, we offer a third, alternative way of explaining predictability. We
consider a simple endowment process, as well as simple agents, but we drop the
representative agent assumption. In particular, we consider agents to interact lo-
cally, in the sense that they take into account the beliefs of their neighbors when
making their consumption decision. Our results demonstrate that an economy
where agents put a large weight on the beliefs of their neighbors is remarkably
successful in explaining several stylized facts from the empirical predictability
literature. In particular, we show that an appropriately calibrated model simulta-
necessarily matches the following empirical facts: (i) From univariate regressions,
future returns seem to be predictable, albeit the statistical significance in not
overwhelming, (ii) In long-horizon regressions the coefficient estimates increase
together with $R^2$s, (iii) The absence of dividend growth predictability is ad-
ditional, even stronger statistical evidence for return predictability, (iv) Regression
coefficients display significant temporal instability, (v) Dividend-yields display
high levels of persistence, and finally (vi) Return predictability is significant even
in the absence of autocorrelations in returns.

Overall, our analysis suggests that ignoring social dynamics means ignoring a sim-
ple, and intuitively appealing way to explain predictability. Our results demon-
strates that incomplete information together with social dynamics can be an
important channel to explain slow-moving, persistent price-dividend ratios.

References


36


