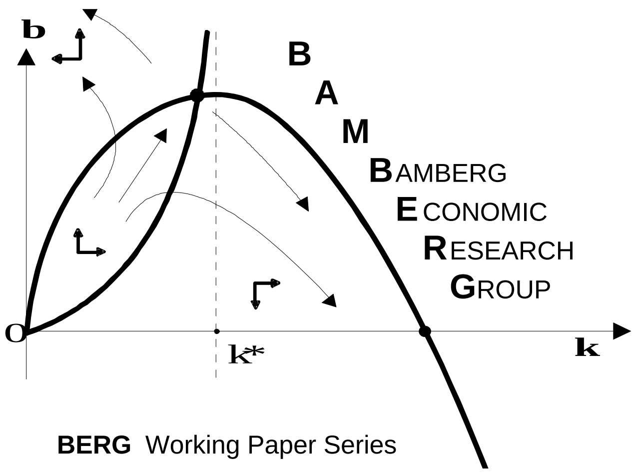
Exploiting ergodicity in forecasts of corporate profitability

Philipp Mundt, Simone Alfarano and Mishael Milakovic

Working Paper No. 147

March 2019



Bamberg Economic Research Group
Bamberg University
Feldkirchenstraße 21
D-96052 Bamberg
Telefax: (0951) 863 5547
Telephone: (0951) 863 2687
felix.stuebben@uni-bamberg.de
http://www.uni-bamberg.de/vwl/forschung/berg/

ISBN 978-3-943153-68-2

Redaktion:
Dr. Felix Stübben*

* felix.stuebben@uni-bamberg.de

Exploiting ergodicity in forecasts of corporate profitability*

Philipp Mundt[†] Simone Alfarano[‡] Mishael Milaković[§]

Abstract

Theory suggests that competition tends to equalize profit rates through the process of capital reallocation, and numerous studies have confirmed that profit rates are indeed persistent and mean-reverting. Recent empirical evidence further shows that fluctuations in the profitability of surviving corporations are well approximated by a stationary Laplace distribution. Here we show that a parsimonious diffusion process of corporate profitability that accounts for all three features of the data achieves better out-of-sample forecasting performance across different time horizons than previously suggested time series and panel data models. As a consequence of replicating the empirical distribution of profit rate fluctuations, the model prescribes a particular strength or speed for the mean-reversion of all profit rates, which leads to superior forecasts of individual time series when we exploit information from the cross-sectional collection of firms. The new model should appeal to managers, analysts, investors and other groups of corporate stakeholders who are interested in accurate forecasts of profitability. To the extent that mean-reversion in profitability is the source of predictable variation in earnings, our approach can also be used in forecasts of earnings and is thus useful for firm valuation.

JEL classifications: C21, C22, C53, L10, D22
Keywords: Return on assets, stochastic differential equation, Fokker-Planck equation, superior predictive ability test, model confidence set.

*We thank Christian Babirat, Sylvain Barde, Thomas Lux, Ilfan Oh, and Christian Proaño for helpful comments and suggestions, and Maximilian Göldl, Nicolas Klas, Michael Neuner, Jan Schulz, and Eva Willer for their able research assistance. P.M. expresses his gratitude to Universitat Jaume I for its hospitality during crucial junctions of this investigation. S.A. gratefully acknowlegdes financial support from Universitat Jaume I, project UJI-B2018-77, and the Generalitat Valenciana, project AICO/2018/036.

[†]Corresponding author. Department of Economics, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany. E-Mail: philipp.mundt@uni-bamberg.de, Phone: (+49) 951 863 2584.

[‡]Department of Economics, University Jaume I, Campus del Riu Sec, 12071 Castellón, Spain.

[§]Department of Economics, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany.

1 Introduction

Accurate forecasts of profitability are relevant for investment decisions and provide valuable information for investors, managers, and other groups of corporate stakeholders. Since refined estimates of profitability improve firm valuation, they are also of interest to financial analysts and traders.[1] Competition tends to equalize profit rates, or in more modern parlor the return on assets (ROA),[2] through the process of capital reallocation that is subject to all types of real frictions. So it is not overly surprising that profit rates are to some extent predictable in the sense that their time series are persistent and mean-reverting (e.g. Nissim and Penman, 2001; Stigler, 1963), which is also true for other accounting ratios such as measures of leverage, liquidity, and operating efficiency (Gallizo et al., 2008). What sets profit rates apart is that their cross-sectional distribution is stationary for surviving corporations and well approximated by a symmetric Laplace density (Mundt et al., 2016). Here we show that a parsimonious diffusion process (or stochastic differential equation) that accounts for all the above statistical regularities in corporate ROA, first suggested by Alfarano et al. (2012), outperforms previously proposed time series and panel data models in terms of their out-of-sample forecasting performance across different time horizons. To the best of our knowledge, the diffusion process is the only model so far that is consistent with both the persistent mean-reversion of individual ROA time series and the cross-sectional distribution of ROA. It is probably best understood as a reduced-form model of economic frictions in the process of capital reallocation, and its distinct feature is that it dictates a precise constraint on the strength of mean-reversion that is derived from the observed cross-sectional distribution.

Our main idea is to exploit the notion of ergodicity in the profitability of surviving corporations. Ergodicity refers to a situation where the unconditional moments of individual time series converge to the moments of the stationary cross-sectional distribution. Put differently, if a system is ergodic then the cross-sectional outcome at a given point in time will convey the same statistical information as the time series of individual destinies; in the natural sciences such a situation is often referred to as a statistical equilibrium (see, e.g., Garibaldi and Scalas, 2010). This is particularly helpful when the number of cross-sectional observations (here several hundred surviving corporations) is larger than the number of observations in the time domain (here several decades of annual data for individual corporations). Our gains in predictive performance therefore originate essentially from exploiting ergodicity both with respect to the law of large numbers, and with respect to the conscientious specification of mean-reversion in individual profit rate series that is prescribed by their cross-sectional distribution. The notion of ergodicity is generally only sensible for surviving entities because in the presence of ruin or corporate death it does not make sense to postulate that cross-sectional or *ensemble* averages are representative of the time series averages of individual entities (see, e.g., Peters and Gell-Mann, 2016; Taleb, 2018). The focus on surviving corporations is, however, less restrictive than it seems at first. After all, if one is not willing to part with concerns of survivorship bias, our results can simply be stated as being conditional on survival. Yet the vast majority of corporate "deaths" are actually caused by transfers of ownership and not by bankruptcy and liquidation, which historically account for a very small fraction of corporate mortal-

[1]The well-known residual income methodology, first suggested by Ohlson (1995) and Feltham and Ohlson (1995), argues that estimates of firm value obtained from (abnormal) earnings are preferable to expected future dividends (see also, for example, the survey by Richardson et al., 2010). Since changes in profitability convey information about changes in earnings (Fama and French, 2000; Freeman et al., 1982), better forecasts of profitability should lead to more accurate forecasts of earnings and thus improve valuation.

[2]In the following we use the terms profitability, profit rate, and return on assets interchangeably.

ity (Daepp et al., 2015). Survivors accordingly carry an enormous amount of incorporated capital through time and also represent macroeconomically crucial "granular" entities in the jargon of Gabaix (2011), making them a worthwhile object of study in their own right.

The reasons for the predictive superiority of our diffusion model are threefold. First, the model is consistent with the empirical Laplace distribution of profit rates, so small deviations around the mean and extreme events occur more often than in models that lead to counterfactual normal distributions. Second, the model is also consistent with the autocorrelation structure of the data that exhibits a particular asymptotic exponential decay, as shown by Mundt et al. (2016). Compared to standard first order autoregressive models, or the Ornstein-Uhlenbeck process as their continuous time analog, our process implies a distinct adjustment towards the average rate of profit that also improves forecasting performance. Third, Mundt et al. (2018) present evidence that individual firm characteristics are almost negligible for the dynamics of profitability once the entity has survived in the market for an extended period of time, pointing to the existence of a common law of motion governing the profitability of long-lived firms that enables us to exploit the ergodic property in the first place.

We employ the test for superior predictive ability (SPA) by Hansen (2005) and the model confidence set (MCS) by Hansen et al. (2011) to evaluate the forecasting performance of our diffusion process against the Ornstein-Uhlenbeck process, more conventional time-series models from the mixed autoregressive integrated moving average (ARIMA) family, and structural and dynamic panel models. In order to distinguish the effect of the diffusion's particular mean-reversion from efficiency gains that originate in the use of panel data, we start out by comparing the forecasting performance of a set of time series models with firm-specific parametrizations that do not make use of cross-sectional information. In this setting our diffusion outperforms alternative time series models such as the Ornstein-Uhlenbeck process, ARIMA models, and the random walk. The finding that our model outperforms these conventional models testifies to the predictability of ROA and the more accurate adjustment mechanism in our model, and it reflects negatively on the so-called persistence of profits literature (see, e.g., Geroski and Jacquemin, 1988; Gschwandtner, 2005; Mueller, 1977, 1990; Waring, 1996), which draws heavily on different types of autoregressive models in their studies. In light of the leptokurtic profit rate distribution, the latter are clearly misspecified and accordingly lead to inferior forecasting performance. Moreover, the Ornstein-Uhlenbeck process as their continuous time analog assumes that the drift towards the average profit rate is stronger the larger the difference from the mean in either direction, as reported by Fama and French (2000) for a mixed sample of firms with different life spans. Our findings indicate that this is not the case for long-lived corporations, testifying to crucial differences in the dynamics of profitability between surviving and shorter-lived corporations.[3] Next we show that the dynamics of profitability are remarkably homogeneous across surviving firms. In particular, our diffusion model exhibits the best forecasting performance when parametrized with estimates of average profitability and dispersion that are obtained from the cross-sectional profit rate distribution. We interpret this as an imprint of ergodicity because the dynamics of all surviving firms follow the same stochastic law that is derived from the cross-sectional return distribution. Such uniformity is not observed for other dimensions of firm behavior, for instance firm size or growth rates therein; Geroski et al. (2003) report that hardly

[3]Even for shorter-lived firms the distribution of profit rates is not Gaussian. Instead, it is more accurately described by an asymmetric exponential power distribution with a "super-Laplacian" left tail that is fatter than that of a Laplace distribution (see, for example, Alfarano et al., 2012; Scharfenaker and Semieniuk, 2016). Fagiolo et al. (2008) also report evidence for "super-Laplacian" distributions in the context of GDP growth rates.

any firm displays systematic growth patterns and that firms do not show any tendency to converge to a common size or to a stationary size distribution, not even in the long-run. In a similar vein, Machado and Mata (2000) cast doubts on the validity of Gibrat's law, which postulates the existence of a common law of motion governing the growth process of all firms. Their study stresses the effect of industry characteristics on growth, which is at odds with the existence of a common mechanism for all companies. Finally we show that alternative models that build on cross-sectional information for parameter estimation, like the structural partial adjustment model introduced by Fama and French (2000) or the dynamic panel models employed by Fairfield et al. (2009), which do not pay attention to the distribution of profit rates, produce on average larger forecast errors than our methodology, especially for longer forecasting horizons where the effect of the correct adjustment mechanism towards average profitability becomes more pronounced.

Despite the appealing statistical properties of the return on assets, the majority of existing papers (mostly in the accounting and finance literature) deals with the modeling and forecasting of earnings as an alternative measure of firm performance. A plethora of contributions from the early literature in this field explores the time-series properties of earnings and concludes that earnings follow a random walk or martingale process, suggesting that the best prediction is simply the last observation (see, e.g., Ball and Watts, 1972; Lintner and Glauber, 1978; Little, 1962). The forecasting capacity of different flavors of time series (mixed autoregressive (integrated) moving average) models on earnings has been investigated by, for example, Albrecht et al. (1977); Brown and Rozeff (1979); Callen et al. (1993); Collins and Hopwood (1980); Foster (1977); Griffin (1977); Lookabill (1976); Watts and Leftwich (1977) while, more recently, Hou et al. (2012) propose a cross-sectional model to forecast the earnings of individual firms. In the research field of corporate profitability, the majority of extant studies employs structural models. Fairfield et al. (1996) analyze the predictive content of several earnings components on the return on equity, such as operating earnings, non-operating earnings and taxes, and special items, finding that disaggregation improves the forecasting accuracy relative to models that use a higher level of aggregation. In a similar vein, Fairfield and Yohn (2001), Soliman (2008), and Bauman (2014) use the DuPont methodology to decompose ROA into the product of asset turnover and profit margin, arguing that changes in asset turnover and profit margin contain information on the change in ROA. More closely related to our approach, several papers employ cross-sectional profitability forecasting models. Fairfield et al. (2009) predict the return on equity and net operating assets by means of dynamic panel models, while Evans et al. (2017), Fama and French (2000), and Allen and Salim (2005) conduct forecasting analysis on profitability using two stage partial adjustment models. Similar to our approach, these models rely on cross-sectional data to predict changes in ROA. Contrary to our model, however, these approaches build on fundamental measures of expected (and thus unobservable) profitability, while our model merely depends on the (observable) history of realized ROA. Yet, the most crucial difference between these models and our methodology is that the former approaches do not take into account the Laplacian nature of profit rate fluctuations.

As the starting point of our methodology is the empirical distribution of firm profit rates from which we subsequently derive a time series model for the dynamics of profitability, our work also relates to the broader body of work focusing on the identification of robust distributional regularities in key economic variables besides firm profit rates. Popular recent examples in the literature include the distribution of firm size (Axtell, 2001; Stanley et al., 1995), firm growth rates (Bottazzi and Secchi, 2003a,b, 2006; Stanley et al., 1996), income (Reed, 2001, 2003; Toda, 2012), consumption (Toda, 2017; Toda and Walsh, 2015, 2017), and GDP growth rates (Fagiolo et al., 2010, 2008). Our approach to

Table 1: Sample composition by sector.

Division	SIC codes	Number of firms
Mining	10-14	19
Construction	15-17	7
Manufacturing	20-39	251
Transportation and public utilities	40-49	79
Wholesale trade	50-51	16
Retail trade	52-59	28
Finance, insurance, and real estate	62-67	39
Services	70-89	36
Total		465

Note: Firms are classified according to the business segment that provided the highest revenue at the end of 2016.

construct a statistical model conditional on such a regularity might, therefore, also prove useful in other applications.

The remainder of this paper is organized as follows. Section 2 describes the data, section 3 outlines the forecasting design, section 4 presents the diffusion model and its competitors, section 5 reports the main results, and section 6 discusses the results and concludes.

2 Data

The dataset for this study is obtained from Thomson Reuters and builds on firms' annual financial statements as reported in the Datastream Worldscope Database. Worldscope includes publicly quoted companies and provides the best coverage and longest history of data for developed markets in North America, where the earliest information is available for 1980. Since this investigation aims at long-lived firms, our focus will be on US companies that were listed on the stock exchange in 1980 and still existed in 2016. This condition is met by 465 entities operating in virtually all sectors of the US economy as shown in Table 1.[4] While it is statistically desirable to select firms with a long history of data in order to maximize the number of observations, the reason for considering surviving corporations goes beyond technical considerations. The selection of long-lived firms is inevitable if we want to exploit the notion of ergodicity. Mean-reversion to a systemic rate of profit implies that (conditional on survival) a firm's abnormally high profit rate will eventually be eroded, while a firm currently operating below the economy-wide average will eventually increase its profitability. This can be thought of as a sort of autopilot mode for surviving firms, a mechanism that is well captured by our reduced-form model. Therefore, the ergodic property of our process, which turns out to be the key element in enhancing the forecasting power, is inherently related to the analysis of surviving corporations. From an economic point of view these firms are an interesting object of study in themselves, as they are often very large along many dimensions of corporate size. The sample used for the present investigation contains more than 200 entities that are listed on the Forbes Fortune 500 list. Given that the gross revenues of the largest 500 US corporations corresponded to approximately 73 percent of US nominal GDP in 2016, their impact on the overall economy is everything but negligible; according to Gabaix's

[4]Following common practice we have merely excluded SIC codes 60 and 61, that is banks, because their total assets are on average one order of magnitude larger than what we observe in any other industry, while their ROA is one order of magnitude smaller than the economy-wide average.

Table 2: Summary statistics for the annual cross-sectional profit rate distribution of 465 surviving corporations.

Year	Mean	SD	Skewness	Kurtosis	Min	Max
1980	0.1404	0.1054	2.5987	21.8771	-0.2581	0.9585
1981	0.1346	0.0948	3.4707	38.9219	-0.1765	1.2094
1982	0.1167	0.0873	0.7863	6.3490	-0.1421	0.5749
1983	0.1180	0.0912	1.3487	11.0316	-0.1915	0.7362
1984	0.1274	0.0783	0.7388	6.4997	-0.1367	0.5687
1985	0.1156	0.0813	0.9776	7.7064	-0.1392	0.6072
1986	0.1022	0.0904	0.1879	7.8897	-0.2771	0.6466
1987	0.1049	0.0851	-0.9026	15.1670	-0.5934	0.5220
1988	0.1083	0.0786	-0.7545	10.7735	-0.4856	0.3465
1989	0.1094	0.0716	0.2402	4.4222	-0.1739	0.3855
1990	0.1028	0.0720	-0.0825	6.3725	-0.3135	0.3931
1991	0.0929	0.0698	0.2087	5.5108	-0.2191	0.4282
1992	0.0948	0.0695	0.4686	5.1601	-0.1655	0.4052
1993	0.0977	0.0672	0.6985	5.7377	-0.1413	0.4220
1994	0.1021	0.0762	-0.3997	11.8677	-0.3714	0.4906
1995	0.1063	0.0751	-0.6523	11.1028	-0.4347	0.4531
1996	0.1066	0.0715	0.0089	6.7887	-0.2435	0.4478
1997	0.1090	0.0750	0.1457	9.0870	-0.2950	0.5896
1998	0.1060	0.0707	0.0597	7.6218	-0.2895	0.4689
1999	0.1032	0.0708	0.2700	5.1047	-0.1854	0.3948
2000	0.1021	0.0805	1.1918	12.8212	-0.2278	0.7243
2001	0.0867	0.0824	2.2753	20.7323	-0.2310	0.8170
2002	0.0807	0.0876	0.3030	30.4418	-0.6976	0.8626
2003	0.0830	0.0822	0.3707	23.2613	-0.6107	0.7044
2004	0.0933	0.0799	2.1754	21.8261	-0.2157	0.8285
2005	0.0957	0.0847	-0.6400	16.1996	-0.5064	0.6448
2006	0.1015	0.0853	-0.0509	19.9470	-0.5756	0.7337
2007	0.0940	0.0935	-1.5832	22.8683	-0.7486	0.6888
2008	0.0854	0.1426	-9.7911	159.1250	-2.2566	0.5060
2009	0.0695	0.0922	-2.1709	23.6643	-0.8252	0.3883
2010	0.0876	0.0901	-2.7693	44.0697	-0.9420	0.6562
2011	0.0945	0.1040	-2.4650	80.4816	-1.2427	1.1218
2012	0.0904	0.1130	1.1152	58.6339	-0.9368	1.3831
2013	0.0915	0.1358	0.0064	120.2388	-1.6702	1.8288
2014	0.0915	0.0946	8.2241	129.9376	-0.1937	1.5689
2015	0.0840	0.1267	-1.4793	88.3218	-1.4924	1.5053
2016	0.0831	0.1070	-4.0337	81.9571	-1.3592	0.9532

"granular hypothesis" their idiosyncratic destinies are responsible for a major fraction of aggregate fluctuations, making them a highly relevant and informative group of firms to study.

We measure profitability in the conventional way, that is in terms of the annual return on assets, computed as the ratio of the flow of operating income to the stock of total assets. Table 2 provides summary statistics for the annual profit rate distributions. They suggest that the cross-sectional profit rate distribution is fairly symmetric around the mean. The year-by-year skewness statistics do not exhibit any clear pattern, indicating that neither negative nor positive skew is a universal feature of the data. Negative realizations of the skewness statistic occur mainly during the last financial and banking crisis, yet it turns out that this is due to extremely few observations. The annual cross-sectional profit rate distributions exhibit considerable excess kurtosis, that is fatter tails than the normal distribution, which is confirmed by the Anscombe and Glynn (1983) test that clearly rejects the null hypothesis of zero excess kurtosis at any level of significance, as shown in Table 9 of Appendix A; various goodness-of-fit tests reject the null hypothesis of normally distributed profit rates for all except one of the 37 annual cross-sectional distributions at the five percent level. Altogether, these empirical observations support the hypothesis of a symmetric leptokurtic profit rate distribution, which has previously been found to approximately follow a double-exponential or Laplace distribution (see, e.g., Al-

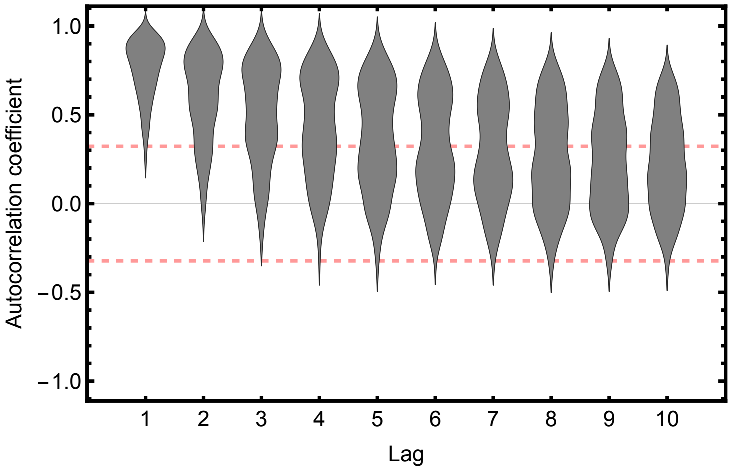


Figure 1: Violin plots of the estimated autocorrelation coefficients for the profit rate time series. Dashed lines correspond to 95 percent asymptotic confidence intervals for zero autocorrelation.

farano and Milaković, 2008; Alfarano et al., 2012; Erlingsson et al., 2012). Turning to the time-series properties, the autocorrelation coefficients that are plotted in Figure 1 suggest that profit rates are positively correlated. This graphical impression is confirmed by the Ljung and Box (1978) and Box and Pierce (1970) tests which both reject the null hypothesis of zero autocorrelations in approximately 90 percent of cases at the five percent level. Accordingly profit rates do not move erratically but appear predictable to some extent due to the rich statistical structure and memory that we find in the data. Finally, we have also tested for stationarity of individual profit rate time-series. Obviously such formal testing is hampered by the rather small number of observations that are available in the time domain and the fact that profit rates are positively autocorrelated, in particular if the speed of adjustment is rather slow. Still, at the 5 percent level we cannot reject the null hypothesis of (second order) stationarity in approximately 70 percent of cases based on the Priestley and Rao (1969) test that considers time-variations in the Fourier spectrum. It is worth noting that a similar frequency is observed for synthetic data of the same length that are simulated from our diffusion process, and that the test easily detects stationarity in simulated data as the length of the time-series is growing larger. We take this to imply that the remaining, approximately 140, profit rate time series are not necessarily non-stationary, and instead attribute this result to the limited power of the test for small samples.[5] In addition both the mean and standard deviation of the annual cross-sectional profit rate distributions, reported in Table 2, exhibit remarkably small fluctuations over time and point towards stationarity in the data from a distributional perspective.

3 Forecasting procedure

To forecast profitability, we employ a rolling window scheme where we estimate a model in-sample and use the fitted model to obtain out-of-sample predictions for forecast horizons

[5]This problem is also relevant for other popular stationarity tests, for instance those which test for unit roots in autoregressions (see, e.g., the discussion in Cochrane, 1991). Notice that the process in eq. (7) is stationary although it exhibits a unit root. Here the stationarity of the process comes from the sign(·) function.

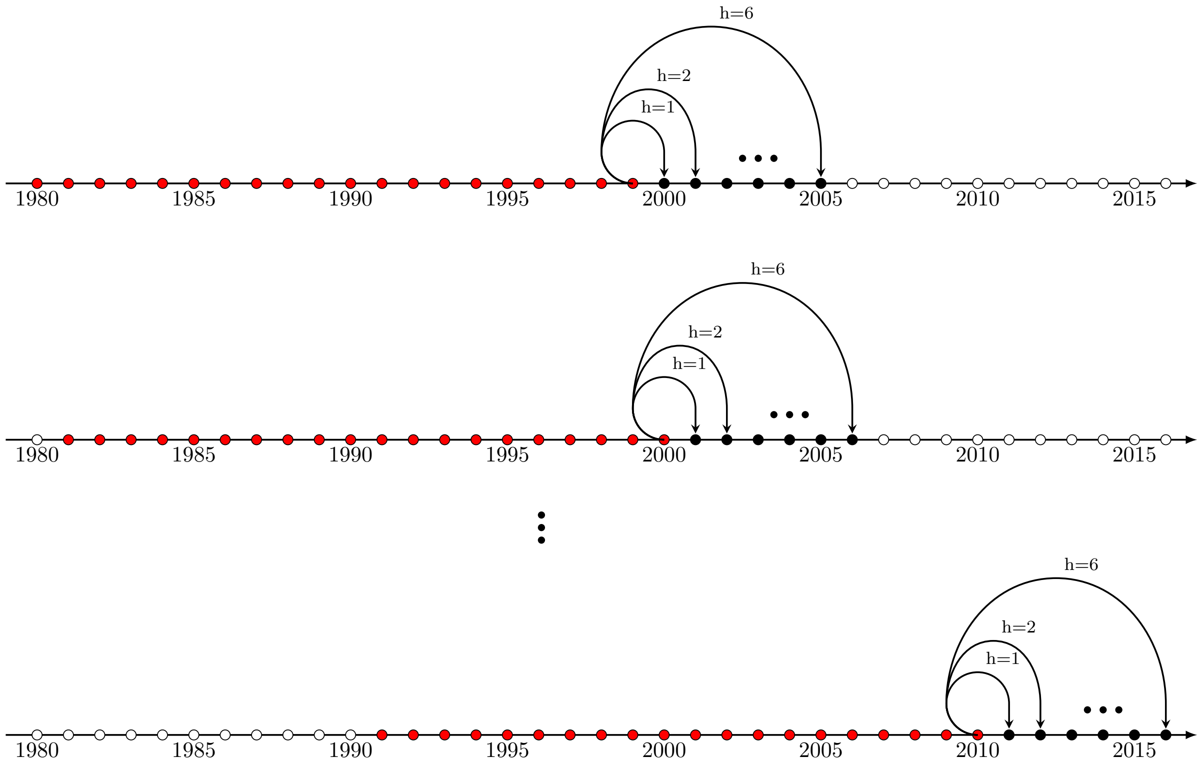


Figure 2: Illustration of h step ahead forecasts with a rolling window. Red circles represent the in-sample period (20 annual observations), while black circles illustrate the testing period (6 annual observations). White circles represent unused data.

of up to 6 years ahead. The rolling window enables us to generate several predictions for each forecast horizon from a single time-series, and it is less sensitive to effects in single years than a fixed scheme because forecast errors are averaged across different time periods. Since the SPA test that we will employ here does not allow parameters to be estimated with a recursive scheme, we do not consider it here either and opt for the rolling window instead.

3.1 Forecast design

To obtain and evaluate the forecasts, we split each time-series into two subsamples. The first one consists of 20 annual observations and serves as an in-sample or training period for parameter estimation, while the following 6 years are used as out-of-sample or testing periods for forecast evaluation. Both the training and testing sample are then rolled forward by adding one more recent observation and dropping one from the beginning of the respective period, so that the number of observations of each of the two subsamples remains constant. Hence, using the available data, we obtain a total of 12 predictions for each forecast horizon $h = 1, 2, \ldots, 6$ years for each firm. Figure 2 provides an illustration of the rolling window scheme.

3.2 Forecast evaluation

The relative out-of-sample forecast accuracy of each model is evaluated by means of the SPA test outlined in Hansen (2005). Unlike the popular Diebold and Mariano (1995) test that examines whether two models exhibit equal predictive accuracy over the entire out-of-sample period, or more refined testing frameworks that consider the time variation in the models' forecasting performance (Giacomini and Rossi, 2010), the SPA test compares the forecasting error of a single (benchmark) model relative to the whole set of competitors. It constitutes a refinement of the so-called *reality check* for data snooping that has been suggested by White (2000), considering a studentized test statistic and a sample-dependent null distribution as modifications. As an additional robustness check, we also consider the model confidence set (MCS) proposed by Hansen et al. (2011) that defines the set of superior models for which the null of equal predictive accuracy cannot be rejected at a given level of significance. Both the SPA and MCS have been designed for applications to time series data. Although neither of them relies on asymptotic theory to obtain the null distribution, the presently available number of observations in the time domain is too small to obtain sufficient variation in the bootstrap samples from which the p-values are derived. Table 10 in appendix B provides Monte Carlo evidence that supports this argument. For simulated time series with the same length as the empirical data ($T = 37$ observations) the results reported in Panel A indicate that neither the SPA nor the MCS test can identify the true data generating process unless we treat the correctly specified model preferentially by estimating its parameters from the cross-sectional distribution. The latter leads to more pronounced differences between the best model and the remaining processes and, therefore, to a higher number of correct identifications. For smaller differences between the models, however, as they would naturally occur in a fair comparison of either time series or cross-sectional models, the tests yield rather inconclusive results. Yet, enlarging the sample size clearly facilitates the identification of the data generating process.

As our investigation builds on a set of panel data for which the number of individuals in the cross-sectional sphere, N, exceeds the number of observations in the time domain, T, by more than one order of magnitude, we attempt to bypass this small sample problem by extending the bootstrap to the cross-sectional dimension as suggested by, for example, Kapetanios (2008). Hence, we use all $N \cdot T$ observations for the computation of average forecast errors and the derivation of the null distribution. The remainder of this section introduces the testing methodology and explains our modification of the tests for the present panel setting, beginning with the SPA test.

Hansen's SPA test asks whether any alternative model is superior to a specific benchmark model in terms of forecast accuracy. To this end, the test considers the null hypothesis that the benchmark is not outperformed by any competing model in terms of expected loss; let x denote the ex-post realized return on assets, and $\hat{x}$ is the corresponding forecast. Then,

$$d_k \equiv L(x, \hat{x}_0) - L(x, \hat{x}_k) \tag{1}$$

quantifies the loss of competing model $k = 1, \ldots, m$ relative to the benchmark model 0, where L denotes some real-valued loss function such as squared or absolute forecast errors.[6] To simplify notation we suppress firm and time subscripts that would appear on all variables of eq. (1). If the benchmark is not outperformed by other models, the average relative loss of each competing model across both the time and the cross-sectional dimension, μ_k, is nonpositive. Hence, the null hypothesis of superior predictive ability of the benchmark model is given by $H_0 : \boldsymbol{\mu} \leq \mathbf{0}$, where $\boldsymbol{\mu} \equiv (\mu_1, \mu_2, \ldots, \mu_m)'$ denotes the vector of (unknown) mean *population* loss differentials. This hypothesis is tested with the test statistic

$$TS^{SPA} \equiv \max_{k=1,\ldots,m} \left[\frac{\sqrt{NT}\bar{d}_1}{\hat{\omega}_1}, \frac{\sqrt{NT}\bar{d}_2}{\hat{\omega}_2}, \ldots, \frac{\sqrt{NT}\bar{d}_m}{\hat{\omega}_m}, 0 \right], \tag{2}$$

[6] In section 5 we report absolute forecast errors, and obtain very similar results for squared losses.

where $T = 12$ denotes the number of observations available in the time domain, $N = 465$ is the number of firms, $\bar{d}_k$ is an estimate of the *sample* mean loss difference, and $\hat{\omega}_k$ denotes the standard error of $\sqrt{NT}\bar{d}_k$ that is employed for studentization. We studentize the mean with

$$\hat{\omega}_k \equiv \sqrt{B^{-1} \sum_{b=1}^{B} \left((NT)^{1/2} \bar{d}^*_{k,b} - (NT)^{1/2} \bar{d}_k \right)^2}, \tag{3}$$

where $d^*_{k,b}$ for $b = 1, 2, \ldots, B$ refers to a bootstrap sample of relative losses. These $B = 5,000$ bootstrap samples have the same size as the empirical data and are obtained by randomly drawing firms with replacement and merging blocks of random length from their time series (see Politis and Romano, 1994). The bootstrap samples also serve to approximate the distribution of the test statistic under the null hypothesis. To this end, we impose the null hypothesis by recentering the bootstrapped performance differentials in $\boldsymbol{\mu}$. Since poor models can bias the test result towards acceptance of the null hypothesis, Hansen (2005) suggests to exclude poor alternatives when simulating the null distribution. Here, we opt for the most liberal test design and exclude all inferior alternatives. Formally, this is achieved by defining the centered variable

$$z_{k,b} \equiv d^*_{k,b} - \max[0, \bar{d}_k]. \tag{4}$$

Denoting the centered bootstrap analog of $\bar{d}_k$ by $\bar{z}_{k,b}$, the distribution of the test statistic under the null is derived from the bootstrapped realizations of the test statistic, and the estimated p-value of the test statistic (2) is given by the percentage of bootstrap statistics that exceeds TS^{SPA}.

A potential shortcoming of our SPA based forecast evaluation is that the present problem lacks a clear benchmark model a priori, so that an overall evaluation of the entire set of models requires sequential testing with alternating benchmarks. [7] In principle, these multiple comparisons of models with the same data face a data snooping bias (see, e.g., Corradi and Distaso, 2011) and thus increase the probability of identifying "superior" models by pure chance. To address this issue, and as an additional robustness check, we also determine the model confidence set (MCS) proposed by Hansen et al. (2011). Similar to a confidence interval for parameter estimates, the MCS represents a set of models which includes the best model with a given probability. An additional advantage of this framework is that the MCS approach may also shed light on the quality of the data in the sense that very informative data may result in an MCS containing only a single model, whereas less informative data will result in a broader MCS.

Starting from an initially complete set of competing models $\mathcal{M}^0$, the idea of the MCS framework is to use an elimination rule to shrink this set to a smaller subset $\mathcal{M} \subset \mathcal{M}^0$ for which the null hypothesis of equal predictive accuracy $H_0 : \mu_{j,k} = 0 \, \forall \, j, k \in \mathcal{M}$, where $\mu_{j,k}$ is the (unknown) mean *population* loss differential between models j and k, cannot be rejected at a certain confidence level. Hence, there is no natural benchmark model against which the performance of all competing models is being tested. Contrary to the SPA test, the MCS considers all possible pairs of models and tests for equivalent forecast performance across these models by means of the test statistic

$$TS^{MCS} \equiv \max_{j,k \in \mathcal{M}} \left| \frac{\bar{d}_{j,k}}{\hat{\omega}_{j,k}} \right|, \tag{5}$$

[7] On the other hand it is hard so see why our model, which is consistent with all three empirical features of ROA, should be less of an *a priori* benchmark model than models that omit at least one statistical feature, typically even violating the empirical ensemble distribution of ROA.

where $\bar{d}_{j,k}$ is the average *sample* loss differential between models j and k. Similarly to the SPA test, the null distribution of this test statistic is obtained from a (cross-sectional) bootstrap procedure because this distribution is (asymptotically) nonstandard. If the test rejects equivalence of the models in the set $\mathcal{M}$, the model with the largest test statistic is discarded from the set and a new test for equal predictive accuracy is run on the remaining models within the set. The algorithm stops when the hypothesis of equivalent forecasting accuracy can no longer be rejected.

4 Competing models

In the following we briefly review our diffusion model, henceforth called AMIK (an acronym formed from the original authors' last names), and its competitors in the forecasting exercise. We include the Ornstein-Uhlenbeck (OU) process as an alternative mean-reverting diffusion, and more standard univariate time-series models from the mixed autoregressive (integrated) moving average varieties, including the special case of a random walk. As alternative cross-sectional models we consider a dynamic panel model and the popular structural partial adjustment model of Fama and French (2000). Details regarding model selection, estimation, and forecasting are provided in each subsection.

4.1 A stochastic model of interacting firms: AMIK diffusion

Alfarano et al. (2012) propose a stochastic model of competitive firms that is analytically tractable in continuous time, arguing that the tendency of competition to equalize profit rates across different economic uses leads to an *equilibrium distribution* of profit rates that characterizes the ensemble of interacting firms. Since the cross-sectional distribution is fairly symmetric and displays significant excess kurtosis, they propose the Laplace distribution as a sensible candidate for distributional fitting

$$f_S(x) = \frac{1}{2\sigma} \exp\left(- \left| \frac{x-m}{\sigma} \right| \right), \tag{6}$$

with location $m \in \mathbb{R}$ and scale (or dispersion) parameter $\sigma > 0$. A series of goodness-of-fit tests in Table 3 suggests that the Laplace distribution is indeed a reasonable approximation of the data inasmuch as the Laplacian null hypothesis cannot be rejected at the 5 percent level for the majority of years. For the remaining cases, the test results indicate departure from the Laplace distribution, which arise from discontinuities in the profit rate distribution around $x \approx 0$ that are well-known in the accounting literature, and most likely have their origin in the management of earnings (see Burgstahler and Dichev, 1997). We fit equation (6) from in-sample ensemble distributions of ROA and report the maximum likelihood (ML) estimates in Table 4, additionally illustrating the empirical densities of in-sample data in Figure 3. Plots of these empirical densities on semi-log scale readily exhibit the linear tent-shape that is characteristic of a Laplace density.

The central idea of the AMIK model is to construct a stochastic differential equation (SDE) for individual profit rate series that has the Laplace distribution in equation (6) as its stationary density. An SDE consists of a drift function and a diffusion function, and its stationary density (provided it exists) depends fundamentally on the ratio of these two functions. This leaves one degree of freedom in the specification of the two functions if we want to impose a certain stationary density, and AMIK parsimoniously assumes a

Table 3: *P*-values of various goodness-of-fit tests for the Laplace distribution.

Year	AD	CVM	KS	KUI	WU2
1980	0.0445	**0.1141**	**0.1299**	0.0034	0.0057
1981	0.0150	0.0418	0.0466	0.0228	0.0056
1982	**0.1207**	**0.1674**	**0.1245**	**0.0610**	**0.1146**
1983	**0.1165**	**0.1747**	**0.1007**	0.0326	**0.0804**
1984	**0.4099**	**0.4995**	**0.4358**	**0.3648**	**0.2809**
1985	**0.1890**	**0.2504**	**0.1287**	**0.1131**	**0.1921**
1986	**0.6543**	**0.6397**	**0.7481**	**0.9573**	**0.9425**
1987	**0.6724**	**0.6600**	**0.5484**	**0.2603**	**0.4444**
1988	**0.1284**	**0.2367**	**0.3772**	**0.2637**	**0.3863**
1989	0.0391	**0.0864**	**0.1056**	**0.1381**	**0.1870**
1990	**0.2523**	**0.3509**	**0.2676**	**0.3345**	**0.4247**
1991	**0.6941**	**0.8145**	**0.7174**	**0.5505**	**0.5958**
1992	**0.2250**	**0.3782**	**0.2916**	**0.4679**	**0.4036**
1993	0.0447	**0.1172**	**0.1190**	**0.3702**	**0.3780**
1994	0.0386	**0.1030**	**0.1480**	**0.1628**	**0.2098**
1995	**0.1393**	**0.2187**	**0.1875**	**0.2668**	**0.2855**
1996	**0.1809**	**0.2476**	**0.2049**	**0.2885**	**0.2607**
1997	**0.0563**	**0.0956**	**0.1511**	**0.1004**	**0.0724**
1998	**0.2050**	**0.3066**	**0.4196**	**0.0910**	**0.1706**
1999	0.0220	**0.0547**	0.0335	0.0116	0.0171
2000	0.0259	**0.0531**	0.0300	**0.1110**	**0.1076**
2001	**0.6540**	**0.8431**	**0.8083**	**0.6508**	**0.8088**
2002	**0.1523**	**0.2830**	**0.0842**	**0.1776**	**0.2438**
2003	**0.0530**	**0.1453**	**0.0645**	**0.1180**	**0.1655**
2004	0.0053	0.0318	0.0052	0.0280	0.0410
2005	0.0254	**0.0911**	**0.0667**	0.0128	0.0150
2006	0.0068	0.0298	0.0154	0.0287	0.0153
2007	0.0189	**0.0591**	0.0121	0.0236	0.0191
2008	**0.1668**	**0.3010**	**0.1995**	**0.1006**	**0.1640**
2009	**0.3376**	**0.4037**	**0.5272**	**0.8286**	**0.6571**
2010	0.0328	**0.1027**	**0.0506**	**0.09879**	**0.0838**
2011	0.0276	**0.0804**	0.0296	0.0120	0.0120
2012	0.0148	**0.0504**	0.0072	0.0040	0.0059
2013	0.0041	0.0226	0.0008	0.0001	0.0002
2014	**0.0864**	**0.1774**	**0.0695**	0.0067	0.0199
2015	0.0080	0.0209	0.0017	0.0011	0.0046
2016	**0.0625**	**0.1057**	0.0073	0.0006	0.0142

Note: We consider the null hypothesis that data were drawn from a Laplace distribution. A small p-value suggests that it is unlikely that the data are Laplacian. Abbreviations refer to AD: Anderson-Darling test, CVM: Cramér-Von Mises test, KS: Kolmogorov-Smirnov test, KUI: Kuiper test, and WU2: Watson U^2 test. P-values greater than 5 percent are shown in boldface. Entries equal to 0.0000 stand for p-values $< 5 \times 10^{-5}$.

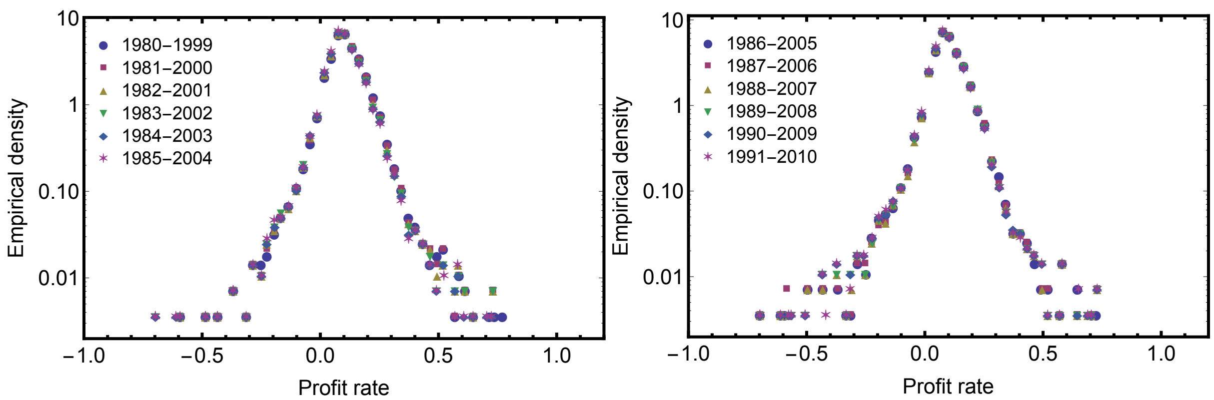


Figure 3: Pooled empirical densities of annual profit rates for 465 long-lived publicly traded US companies during different in-sample periods.

Table 4: Estimated in-sample parameters of the fitted Laplace distribution.

Period	$\hat{m}$	SE($\hat{m}$)	$\hat{\sigma}$	SE($\hat{\sigma}$)
1980-1999	0.1031	0.0008	0.0567	0.0006
1981-2000	0.1015	0.0007	0.0559	0.0006
1982-2001	0.0998	0.0007	0.0554	0.0006
1983-2002	0.0982	0.0007	0.0550	0.0006
1984-2003	0.0963	0.0008	0.0545	0.0006
1985-2004	0.0942	0.0008	0.0541	0.0006
1986-2005	0.0932	0.0007	0.0539	0.0006
1987-2006	0.0927	0.0007	0.0535	0.0006
1988-2007	0.0921	0.0006	0.0535	0.0006
1989-2008	0.0914	0.0006	0.0540	0.0007
1990-2009	0.0901	0.0006	0.0544	0.0007
1991-2010	0.0894	0.0006	0.0545	0.0007

Note: The table shows the estimated location and dispersion parameters of a Laplace distribution for different samples of training data and the standard errors of the parameter estimates.

constant diffusion function to derive a drift function that will lead to a stationary Laplace density,[8] resulting in the mean-reverting SDE

$$dX_t = -\frac{D}{2\sigma}\text{sign}(X_t - m)dt + \sqrt{D}\,dW_t. \tag{7}$$

The SDE in eq. (7) defines a regular diffusion on the real line around the measure of central tendency $m \in \mathbb{R}$ with dispersion $\sigma > 0$; the parameter $D > 0$ is called diffusion coefficient and dW_t are standard Wiener increments. In the following, we denote the set of these three parameters by $\boldsymbol{\theta} = \{m, \sigma, D\}$.

The mean-reverting drift function in the deterministic first term of the SDE reflects the systematic tendency for competition to equalize profit rates across firms, while the diffusion function in the second term incorporates idiosyncratic random shocks to profitability with mean zero and variance D. A peculiar feature of the AMIK drift is that current realizations of the profit rate solely determine the sign of the drift and not its strength, which in absolute terms is equal to $D/(2\sigma)$. As we shall see momentarily, this constitutes a key difference to the competing OU process for which the strength of the drift depends linearly on current deviations from the average profit rate m. An important conceptual aspect of the AMIK diffusion concerns the fact that the level of idiosyncratic noise D, which is absent in the stationary distribution (6), shows up in both the deterministic drift and the random diffusion function. In other words, the AMIK diffusion decomposes the metaphor of competition into the contemporaneous presence of idiosyncratic fluctuations and a systematic tendency towards profit rate equalization, which jointly give rise to a stationary distribution that is Laplace under the particular drift in eq. (7). From the viewpoint of ergodicity, it is important to realize that the distributional regularity in eq. (6) refers to cross-sectional data, as for instance in Figure 3, while the dynamic law in eq. (7) refers to individual time-series. If the system is ergodic, cross-sectional and time-series properties coincide and the parameter values of the ensemble distribution carry over to the time series of surviving firms, leaving D as the only source of idiosyncratic variation in the time evolution of their profitability.

[8] Alfarano et al. (2012) consider the more general problem of constructing an SDE that has the exponential power distribution as its limiting density, so the Laplace diffusion in eq. (7) is actually just a special case of their model.

To estimate the diffusion process, we proceed as follows. Since equation (7) defines a Markov process with continuous trajectories, its transition probabilities obey the Fokker-Planck equation (see, e.g., Gardiner, 2009; Risken, 1996)

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}(A(x;\boldsymbol{\theta})p(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(B(x;\boldsymbol{\theta})p(x,t)), \tag{8}$$

where $A(x;\boldsymbol{\theta})$ and $B(x;\boldsymbol{\theta})$ are the drift and diffusion functions of the underlying process (in the AMIK case, $A = -D\,\mathrm{sign}(X_t - m)/(2\sigma)$ and $B = \sqrt{D}$), and $p(x,t) \equiv f(x,t|x_0,0)$ denotes the conditional probability density for a transition from some initial state x_0 at time 0 to state x at time t. The Fokker-Planck equation is a deterministic second-order partial differential equation that is employed for ML estimation of the parameters and for obtaining the predictive density. Considering the initial condition of a unit probability mass in point x_0, Toda (2012) shows that the solution to the Fokker-Planck equation for the AMIK diffusion is

$$\begin{aligned} f(x,t|x_0,0) = {} & \frac{1}{\sqrt{2\pi Dt}} \cdot \exp\left\{-\frac{(x-x_0)^2}{2Dt} - \frac{1}{2\sigma}(|x-m| - |x_0-m|) - \frac{Dt}{8\sigma^2}\right\} \\ & + \frac{1}{2\sigma}\exp\left\{-\frac{1}{\sigma}|x-m|\right\}\Phi\left(-\frac{|x-m| + |x_0-m| - \frac{Dt}{2\sigma}}{\sqrt{Dt}}\right), \end{aligned} \tag{9}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Since there is no previous observation for x_0, we follow standard procedure and evaluate this component using the unconditional probability density $f_S(x)$ (see, e.g., Ghongadze and Lux, 2012; Lux, 2009).[9] The log-likelihood of observations then amounts to

$$\log \mathcal{L}(\boldsymbol{\theta}) = \log f_S(x_0;\boldsymbol{\theta}) + \sum_{t=0}^{T-1} \log f(x_{t+1}|x_t;\boldsymbol{\theta}), \tag{10}$$

and we maximize it numerically to obtain the parameter estimates $\hat{\boldsymbol{\theta}} = \{\hat{\sigma}, \hat{D}\}$. Notice that the location parameter m is not estimated from eq. (10) because the corresponding ML estimator $\hat{m}_{ML}$ violates standard regularity conditions due to the singularity of the AMIK diffusion in m, implying that $\hat{m}_{ML}$ is not the minimum variance unbiased estimator of the location parameter (see Kotz et al. (2001), p. 64ff., for a discussion of this issue in case of Laplace distributed random variables, or Mittelhammer (2013), chap. 7, in a more general context). We have investigated in Monte Carlo studies that there are alternative asymptotically unbiased estimators of m with smaller variances than the ML estimator, in particular in small samples, and hence opt for the trimmed mean instead of the ML estimator.[10] Given this estimate of average profitability, eq. (10) is maximized with respect to the remaining two parameters that capture the dispersion of the process and the level of idiosyncratic noise, which together determine the speed of mean-reversion in the AMIK diffusion.

The finding that the cross-sectional profit rate distribution is stationary Laplacian suggests that the dynamic law in eq. (7) applies to all survivors and thus prescribes a common measure of average profitability and dispersion. To check the validity of this hypothesis, we distinguish two alternative model specifications, summarized in Table 5. While AMIK TS estimates all three parameters $\{m, \sigma, D\}$ individually for each firm, in the AMIK CS specification we impose the restriction that both the location and scale

[9] Alternatively, one could simply drop the first observation. As we have only 20 in-sample observations, we decided against this practice and instead follow the procedure described in the main text.

[10] This material is available upon request.

Table 5: Alternative specifications of the AMIK model.

Specification	Cross-sectional parameters	Idiosyncratic parameter(s)
AMIK TS		m, σ, D
AMIK CS	m, σ	D

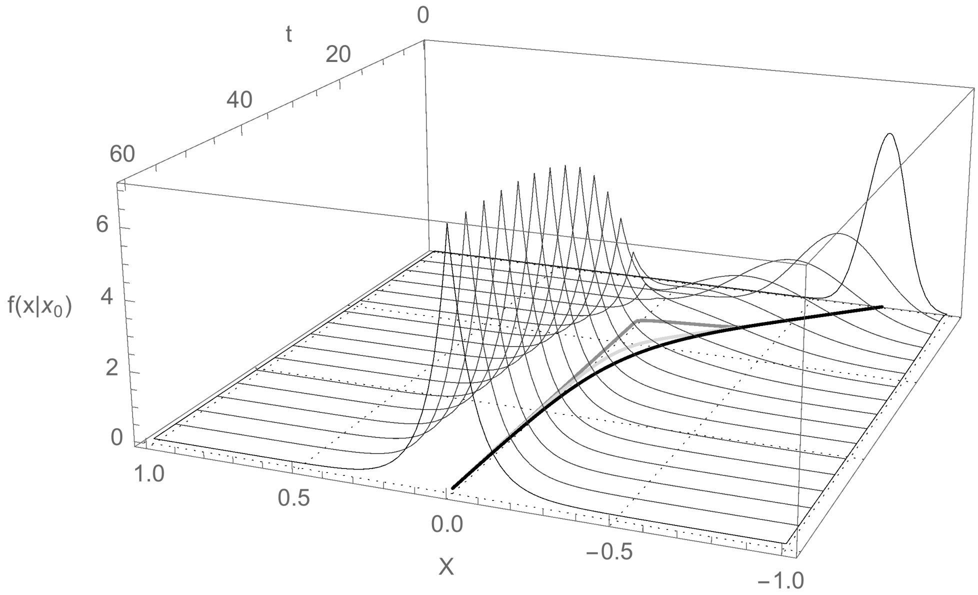


Figure 4: Evolution of the transient density of the AMIK diffusion conditional on the starting value $x_0 = -0.8$. The diffusion coefficient equals $D = 0.005$, and the standard deviation is $\sqrt{2}\sigma = 0.1$. As time increases the mode shifts toward $m = 0$. The three bold lines represent the mean (black), median (gray), and mode (light gray) of the transient density.

parameter are identical across all firms and correspond to the phenomenological values $\hat{m}$ and $\hat{\sigma}$ of the cross-sectional profit rate distribution, implying that the profitability of all survivors reverts to the same average and exhibits the same dispersion under the ergodic hypothesis that is central to the AMIK model.

In addition to parameter estimation, we can also employ the solution of the Fokker-Planck equation as our predictive distribution. As illustrated in Figure 4, we observe an asymmetric conditional probability density in the transient regime of the process that converges to the symmetric Laplace with unconditional mean m and long-term dispersion σ for $t \to \infty$. Since forecast horizons of up to 6 years are typically not long enough to reach this stationary distribution, the mean, mode, or median of the transient density become natural candidates for the forecast. Here we focus on the mean prediction because it turns out to provide the best forecasting performance; the expected value is then determined by numerically solving the integral

$$E_t[X|x_0] = \int_{-\infty}^{\infty} x f(x, t|x_0, 0) dx. \tag{11}$$

4.2 Ornstein-Uhlenbeck process

An alternative mean-reverting diffusion process that is fully analytically tractable is the prominent model of Uhlenbeck and Ornstein (1930)

$$dX_t = \frac{D}{2\sigma^2}(m - X_t)dt + \sqrt{D} dW_t, \tag{12}$$

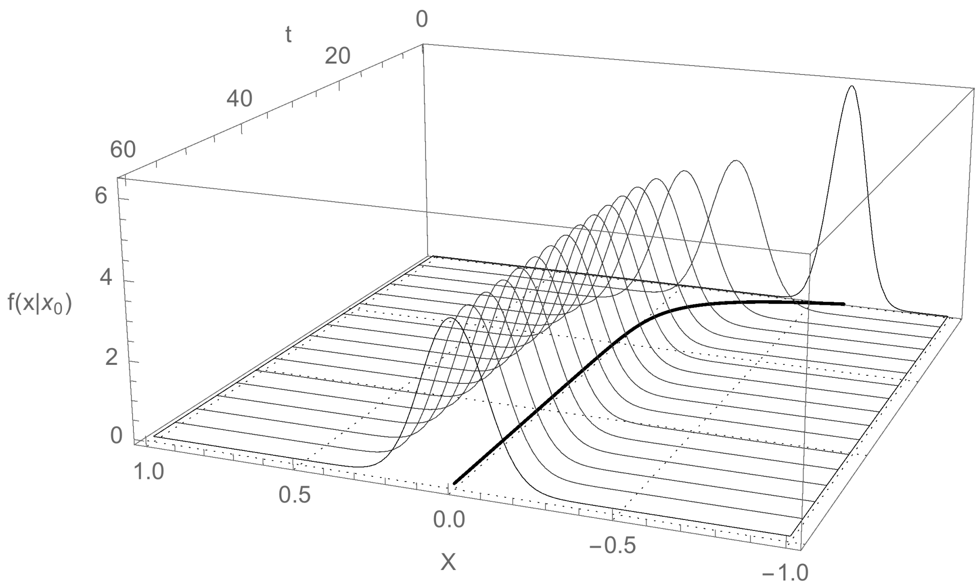


Figure 5: Evolution of the transient density of the OU process conditional on the starting value $x_0 = -0.8$. The diffusion coefficient equals $D = 0.005$ and the standard deviation is $\sigma = 0.1$. As time increases the mode shifts towards $m = 0$. The bold black line represents the mean prediction.

that is the continuous time analog of a stationary first order autoregressive process.[11] The latter has frequently been used as a model of profitability in the so-called persistence of profit literature (e.g. Geroski and Jacquemin, 1988; Goddard and Wilson, 1999; Gschwandtner, 2005; Mueller, 1986, 1990). Contrary to AMIK, the OU process has a drift term that depends linearly on the deviation from the unconditional mean m.[12] Put differently, the greater the distance of the actual realization from the long-term average m, the stronger the drift that pulls the profit rate back to its unconditional mean. Extreme profit rate realizations thus occur less frequently than in the AMIK model, leading to a more platykurtic distribution than the Laplace. The solution to the Fokker-Planck equation (8) for the OU process (12) is also analytically tractable and reads

$$\begin{aligned} f(x,t|x_0,0) = & \frac{1}{\sqrt{2\pi\sigma^2\left(1-\exp\left(-\frac{Dt}{\sigma^2}\right)\right)}} \\ & \times \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{\left((x-m)-(x_0-m)\exp\left(-\frac{Dt}{2\sigma^2}\right)\right)^2}{1-\exp\left(-\frac{Dt}{\sigma^2}\right)}\right)\right\} \end{aligned} \tag{13}$$

which is Gaussian for all t, with time-dependent first and second moments (see, e.g., Gardiner, 2009, p. 128). Figure 5 plots the transient density as a function of time and illustrates how the conditional probability density converges to a stationary normal distribution

$$f_S(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \tag{14}$$

[11] Notice that the generic dispersion parameter σ refers to different dispersion measures in equations (7) and (12). In the Laplacian AMIK model, σ stands for the mean absolute deviation, while it denotes the standard deviation in the Gaussian OU process; we can relate both measures because a Laplace distributed random variable with scale parameter σ has a standard deviation of $\sqrt{2}\sigma$.

[12] Another diffusion process that exhibits mean-reverting behavior is the model by Cox et al. (1985) that is frequently used for the modeling of interest rates. In their model the conditional dispersion of random innovations in the diffusion function prevents negative realizations of the process. Owing to the fact that a zero lower bound is nonsensical for ROA, we do not consider their model here.

as $t \to \infty$, with unconditional mean m and variance σ^2. The OU process is estimated with the ML method using equations (10), (13) and (14). For a fair comparison with the AMIK model we also distinguish between two specifications of the process, one where the three parameters are estimated from the profit rate time series of an individual company (OU TS), and one cross-sectional model where average profitability and dispersion are estimated from pooled data using a Gaussian density (OU CS). In each specification the forecast is the conditional mean of the analytical transient density

$$E_t[X|x_0] = x_0 \exp\left(-\frac{Dt}{2\sigma^2}\right) + m\left(1 - \exp\left(-\frac{Dt}{2\sigma^2}\right)\right). \tag{15}$$

4.3 ARIMA-type time series models

As additional candidates we consider time-series models from the mixed autoregressive and moving average varieties. The ARMA(p, q) model reads

$$(1 - \sum_{i=1}^{p} \lambda_i B^i) X_t = c + (1 + \sum_{j=1}^{q} \psi_j B^j)\varepsilon_t, \tag{16}$$

where B represents the backshift operator, $\Lambda(B) = 1 - \lambda_1 B - \lambda_2 B^2 - \cdots - \lambda_p B^p$ and $\Psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \cdots + \psi_q B^q$ are the autoregressive and moving average polynomials, respectively, and ε_t denotes a white noise series with $E[\varepsilon_t] = 0$, $E[\varepsilon_t^2] = \sigma^2$ and $E[\varepsilon_t \varepsilon_\tau] = 0$ for $t \neq \tau$. The constant c is included to capture a possibly non-zero mean of the process. We estimate the $p + q + 2$ parameters in $\boldsymbol{\Xi} = (c, \lambda_1, \lambda_2, \ldots, \lambda_p, \psi_1, \psi_2, \ldots, \psi_q, \sigma^2)'$ for $1 \leq p, q \leq 5$ via ML. From the set of estimated models within this range, we then choose the specification that minimizes the Schwarz information criterion, which tends to select more parsimonious models than the Akaike information criterion. To forecast with ARMA models we follow the standard procedure for obtaining linear forecasts as outlined, for example, in Diebold (2006). As a special case of ARMA(1,0) with unit root, we also consider the random walk (RW) without drift. Obviously RW requires no parameter estimation as the best prediction is simply the last observation. Finally, we consider mixed autoregressive integrated moving average models

$$(1 - \sum_{i=1}^{p} \lambda_i B^i)(1 - B)^d X_t = (1 + \sum_{j=1}^{q} \psi_i B^j)\varepsilon_t \tag{17}$$

to refute potential concerns regarding non-stationarity of the data.

4.4 Structural partial adjustment model

Another competitor is the partial adjustment model (PAM) proposed by Fama and French (2000) that builds on a two step cross-sectional regression. The first regression is a structural model that derives an estimate of expected profitability from a set of accounting variables, while the second regression predicts the change of profitability. The adjustment is modeled as

$$X_{i,t+1} - X_{i,t} = \alpha_t + \beta_t (X_{i,t} - X_{i,t}^e) + \gamma_t (X_{i,t} - X_{i,t-1}) + \epsilon_{i,t+1}, \tag{18}$$

where $X_{i,t} - X_i^e$ represents the deviation of firm i's current profit rate from its expected value. Idiosyncratic estimates of expected profitability are obtained from a structural specification using fitted values from the regression

$$X_{i,t}^e = \delta_t + \zeta_t V_{i,t}/TA_{i,t} + \eta_t DD_{i,t} + \theta_t D_{i,t}/BE_{i,t} + \varepsilon_{i,t} \tag{19}$$

that relates expected profitability to the ratio of the market value of equity and debt to the book value of total assets, V/TA (essentially a proxy for Tobin's q), to the ratio of paid dividends to the book value of common equity, D/BE, and to a dummy DD that equals 0 for dividend paying firms and 1 otherwise. The rationale for estimating this structural relationship is that dividends supposedly contain information about expected future earnings, and that dividend payers tend to be more profitable than non-payers (Fama and French, 2001, 2002). Finally, the market-to-book value of assets is included to account for variation in expected profitability that is not captured by dividends.

To obtain forecasts from eq. (18) and (19), we follow the same procedure as Fama and French (2000) and compute the average intercepts and slopes from the cross-sectional regressions across all 20 years of in-sample data. From these we obtain estimates of expected profitability and the change in profitability. Eq. (18) implies that forecasts of expected profitability are necessary to predict the change in profitability for more than one step ahead, so we assume that the expected future (out-of-sample) profitability of firm i is equal to the average value of expected profitability across all 20 years of the training period.

PAM demands substantially more data to model and forecast changes in profitability compared to its competitors, which is not only problematic from the viewpoint of Occam's razor but also leads to more practical problems when the required data are (temporarily) unavailable. Indeed we face a situation where expected profitability cannot be computed for some corporations because of missing values in dividend payments or other quantities in particular years. In these cases we extrapolate the expected profit rate by computing the time average of the available estimates from other in-sample periods. In the rare cases where we face missing values for all training samples, we approximate expected profitability with the sample mean profit rate.[13] Finally, forecasts of profitability are obtained from eq. (18) by adding the previous realization of the profit rate to the predicted change in that variable.

4.5 Non-linear partial adjustment models

Fama and French (2000) and Fairfield et al. (2009) consider models that assume an asymmetry in the mean-reverting behavior of firms with low and high profitability. They argue that mean reversion is faster when profitability is below its mean or, equivalently, that low profitability is less persistent than high profitability. To account for this supposed asymmetry, consider the economy-wide non-linear partial adjustment model (NLPAM EW)

$$X_{i,t} = \alpha_t + \beta_t X_{i,t-1} + \gamma_t D_t X_{i,t-1} + \epsilon_{i,t}, \tag{20}$$

where D_t is an indicator variable that equals one if the deviation of profitability from its median was negative in the previous period $t-1$, and zero otherwise. Allowing for sector dependencies of the parameters α and β, we can also check whether profitability converges to industry-specific (IS) rather than economy-wide (EW) averages, a question that has been addressed in the study by Fairfield et al. (2009). Formally, the industry-specific model requires the modification

$$X_{i,t} = \alpha_{j,t} + \beta_{j,t} X_{i,t-1} + \gamma_{j,t} D_t X_{i,t-1} + \epsilon_{i,t}, \tag{21}$$

[13] An alternative would be to simply exclude these firms from the analysis. As we aim to maximize sample size and since profit rate data are none the less available for these companies, we decided against this practice.

where the index j runs over the business divisions in Table 1. Henceforth, we will refer to this sectoral model as NLPAM IS.[14] The estimation strategy is geared to that of Fama and French (2000) and Fairfield et al. (2009), that is for each year we estimate cross-sectional regressions using pooled annual data and then compute the time averages of the estimated parameters. NLPAM EW assumes that the profitability of all firms reverts to the same average and thus considers the full cross-section of firms, while the industry-specific model NLPAM IS considers only those firms operating in the same industry.

5 Results

Before turning to the comparison of forecasting performances, we start out by discussing the estimation results for the competing model specifications, comprehensively reported in Tables 11-18 of Appendix C.[15] Considering the model specification AMIK TS, we can see that the estimates of the location parameter are in line with the respective ensemble values, while the estimates of the dispersion parameter are notably smaller than the ensemble dispersion. The latter occurs because σ measures the dispersion of profit rates around the long-run mean, implying that the time series estimates of average profitability (which tend to be subject to overfitting in small samples) lead to smaller estimates of dispersion on average.[16] Results for OU and AMIK TS are similar with respect to the magnitude of the estimated parameters, yet the OU process predicts faster mean-reversion due to the appearance of the variance term σ^2 in the drift function.

Considering the class of ARMA(p, q) models, the parsimonious ARMA(1,1) specification comes out as the most favored model based on the Schwarz criterion and is therefore chosen for forecasting. The estimated autoregressive parameters are smaller than unity in absolute value in 431-446 out of 465 cases (depending on the period under consideration), indicating stationarity for the vast majority of fitted processes. As an additional robustness check, we also consider integrated autoregressive moving average models: for $1 \leq p, q \leq 5$ and $d \geq 1$, ARIMA(1,1,1) compares most favorably based on the Schwarz criterion, so we fit ARMA(1,1) to the first difference of the original series which is then used for forecasting.

Estimation results for the cross-section partial adjustment model are largely in line with those reported by Fama and French (2000) for the US and Allen and Salim (2005) for the UK. We can confirm that Tobin's q is positively correlated with the profit rate. Although the significant dividend dummy lends support to Fama and French's conjecture that the relation between expected profitability and dividends might be non-linear, the aggregate effect of dividends on returns appears to be rather fragile as we do not observe a significant effect of the dividend payout ratio on expected profitability in every period under consideration. The second regression regarding the change in profitability shows that the partial adjustment term is significantly negative, as is to be expected from the mean-reverting property of the ROA time series. Our estimate of the average rate of mean-reversion is about 18 percent per year, which is smaller than the 38 percent reported by Fama and French for US data, and closer to the 25 percent reported by Allen

[14] The two models in eq. (20) and (21) are almost identical to those considered in Fairfield et al. (2009). The only difference is that they consider the predicted sales growth rate as an additional explanatory variable for profitability.

[15] For time series models we report summary statistics of the parameter estimates.

[16] Notice that differences between time series and cross-sectional estimates of the parameters do not necessarily contradict the ergodic hypothesis because a slow adjustment speed, measured by the (idiosyncratic) diffusion coefficient, may lead to time series estimates of the location parameter that differ substantially from the cross-sectional average in small samples. As described in the main text, this will also affect the estimates of σ because the latter depends on m.

and Salim for the UK. The coefficient for the lagged change in profitability is close to zero and not statistically significant for some training samples. Thus, unlike Fama and French, we do not find robust evidence for autocorrelations in profitability that go beyond the adjustment towards average profitability. We conjecture that this is a consequence of considering surviving firms for which the mean-reversion towards the average is the only systematic effect governing the change in profitability. Since we suspect that the variability in expected profitability and the change in profitability are mainly explained by Tobin's q and the deviation from expected profitability, we also estimate simpler models that omit dividends and the lagged change in ROA as explanatory variables in the respective regression equations. We indeed find that the performance of Fama and French's model hinges crucially on Tobin's q and the partial adjustment term because the (adjusted) R-squared falls by merely seven to fifteen percent when the two regressors are excluded (see Table 17 in the appendix). In some training samples, the Bayesian information criterion even indicates weak superiority of the more parsimonious regression model for the change in profitability. We take this to imply that the original formulation of the partial adjustment model is most likely overfitted, and that its explanatory power depends mainly on the strong correlation between book rates of return and market valuation.

Finally, for the non-linear partial adjustment models NLPAM EW and NLPAM IS, we observe that the estimates of the partial adjustment coefficient are smaller than unity and highly statistically significant across all periods and model specifications. For the economy-wide model, the slope coefficient lies in the range $0.8 - 0.85$, while we observe slightly more variation in this parameter for the industry-specific model, suggesting the presence of differences in the adjustment speed across industries. Notice, again, that such heterogeneity is also consistent with AMIK CS, where the adjustment speed is captured by the diffusion coefficient D, which is the only idiosyncratic parameter in the AMIK CS model. We obtain less stringent results for the low profitability dummy in NLPAM EW and NLPAM IS specifications, whose significance depends on both the period and the industry under consideration. Those coefficients that are statistically significant are consistently negative, suggesting that low profitability might be (asymmetrically) less persistent than high profitability, as predicted by Fama and French (2000) and Fairfield et al. (2009), which is at odds with the constant (and thus symmetric) nature of the drift in the AMIK process that is derived from the (symmetric) Laplacian ensemble distribution.

To sum up, the competing models exhibit substantial heterogeneity in their estimated parameterizations and pre-analytical visions to warrant a comparison of their forecast performance.

5.1 Time series models versus the naïve forecast

Our assessment of predictive accuracy starts out by comparing the AMIK TS diffusion, parametrized from firm-specific estimates of average profitability and dispersion, to alternative time series models. Hence we disregard potential improvements in AMIK that arise from the use of cross-sectional information at first, and initially focus on the adjustment mechanism towards average profitability. The results in Panel A of Table 6, which summarize the mean absolute forecast error for the different models, suggest that the AMIK model yields the most accurate forecasts for medium to long-run horizons.[17] Yet for pre-

[17] Since the transient density of the AMIK diffusion becomes asymmetric for some t, we have also experimented with alternatives to the mean prediction. It turns out that both the median and the mode of the conditional probability density function are dominated by the expected value, i.e. their (unreported) forecasting results are clearly inferior to the mean prediction. The failure of the mode arises from the abrupt change of the transient density's maximum that can be observed in Figure 4.

Table 6: Forecast errors and statistical tests of relative forecast accuracy for alternative time series models.

Horizon h	AMIK	OU	ARMA	ARIMA	RW
Panel A: Mean absolute forecast errors					
1	0.0176	0.0175	$\gg 1$	0.3350	**0.0161**
2	0.0236	0.0238	$\gg 1$	$\gg 1$	**0.0231**
3	**0.0265**	0.0275	$\gg 1$	$\gg 1$	0.0272
4	**0.0287**	0.0296	$\gg 1$	$\gg 1$	0.0292
5	**0.0307**	0.0314	$\gg 1$	$\gg 1$	0.0321
6	**0.0319**	0.0329	$\gg 1$	$\gg 1$	0.0327
Panel B: *P*-values of the superior predictive ability test					
1	0.0000	0.0000	0.0000	0.0000	**1.0000**
2	**0.1650**	0.0000	0.0000	0.0000	**1.0000**
3	**1.0000**	0.0000	0.0000	0.0000	**0.1660**
4	**1.0000**	0.0000	0.0000	0.0000	**0.1010**
5	**1.0000**	0.0000	0.0000	0.0000	0.0032
6	**1.0000**	0.0000	0.0000	0.0000	0.0376
Panel C: *P*-values of the model confidence set					
1	**0.2932**	**0.2932**	**0.2932**	**0.2932**	**1.0000**
2	**1.0000**	**0.1260**	0.0560	0.0560	0.0560
3	**1.0000**	0.0216	0.0006	0.0006	0.0006
4	**1.0000**	0.0236	0.0012	0.0012	0.0012
5	**1.0000**	0.0456	0.0006	0.0006	0.0006
6	**1.0000**	0.0038	0.0002	0.0002	0.0002

Note: Panel A shows the mean absolute forecast errors (MAE) for pooled data. Bold values indicate the best model(s) in terms of the smallest average loss. Entries $\gg 1$ imply that the average forecast error exceeds unity. Panel B reports the *p*-values of the test for superior predictive ability by Hansen (2005). The null hypothesis is that the respective model (i.e. each respective column) is not outperformed as a benchmark by any other of the four remaining models in terms of MAE. Panel C presents the MCS *p*-value for each competing model that expresses the probability of this particular model belonging to the model confidence set introduced in Hansen et al. (2011). A low *p*-value indicates rejection of the null hypothesis at the pertinent confidence level. *P*-values larger than 10 percent are shown in boldface. In all three panels, AMIK and OU refer to the time series (TS) specification of the respective process where all parameters are estimated from the profit rate time series of each firm.

dictions of up to two years ahead the naïve RW performs better than the mean-reverting AMIK process and the OU model. We attribute this result to the high persistence of firm profitability that causes the last observation to be a good short-run predictor, and to the lack of parameter uncertainty in the random walk model. The AMIK process, however, compares favorably to the random walk for longer time horizons. In addition, we can also observe superior predictive accuracy of AMIK relative to OU across all forecasting horizons, and the size of the advantage increases with the length of the forecast horizon for which our (AMIK) specification of the adjustment mechanism becomes increasingly important. As argued before, this adjustment process is more accurately described by the AMIK diffusion because it is consistent with the empirical profit rate distribution, whereas the OU process is obviously counterfactual in light of its stationary Gaussian distribution. Models from the family of autoregressive (integrated) moving average processes are distinctly dominated by AMIK, OU, and RW and there is no indication that differencing improves the forecasting accuracy since ARIMA performs worse than ARMA. The average forecast errors of these models exceed those of AMIK by several orders of magnitude, which turns out to be caused by several firms whose estimated autoregressive parameters are greater than unity in absolute value. ARMA and ARIMA models still

Table 7: Incremental improvement in forecast accuracy.

Horizon h	Improvement	P-value	Improvement	P-value
	AMIK CS vs. AMIK TS		AMIK CS vs. RW	
1	0.0004	0.0000	0.0002	**0.3431**
2	0.0006	0.0000	0.0009	0.0000
3	0.0009	0.0000	0.0017	0.0000
4	0.0011	0.0000	0.0020	0.0000
5	0.0008	0.0000	0.0026	0.0000
6	0.0008	0.0000	0.0029	0.0000

Note: We report the mean improvement in forecast accuracy as measured through a pairwise comparison of absolute forecast errors between the two competing models. A positive (negative) value indicates that the first mentioned model is more (less) accurate than the second model. Tests of means are based on a t test. The null hypothesis is that the average difference is equal to zero. P-values greater than 10 percent are shown in boldface and entries equal to 0.0000 imply p-values $< 5 \times 10^{-5}$.

remain inferior to AMIK, OU, and RW if these firms are excluded in the computation of cross-sectional average forecast errors.

The SPA test results, summarized in Panel B of Table 6, support the view that the AMIK process is more accurate than the random walk or alternative mean-reverting models for longer forecast horizons. For $h \geq 3$, AMIK achieves the maximum $p-$value and we cannot reject the null hypothesis that it is superior to its competitors. On the other hand, for $h \leq 4$ years the p-values of the random walk are not small enough to reject the null at the usual confidence levels either, indicating that the gains in forecast accuracy of AMIK, at least when parametrized with firm-specific parameters for average profitability and the magnitude of fluctuations, are relatively minor for short time horizons. For $h \geq 5$ years, the AMIK process evidently outperforms the different flavors of (mixed) autoregressive and random walk models that have been widely applied in the field of corporate profitability.

To investigate the impact of a potential sequential testing bias on our results, and as an additional robustness check, we also consider the model confidence set as an alternative methodology to compare forecast accuracy in Panel C of Table 6. Overall, the findings corroborate the impression from the SPA test and support the validity of our results for different methods of forecast evaluation.

5.2 Exploiting ergodicity

AMIK CS assumes that the profit rates of all (surviving) firms obey the same law in eq. (7) and thus have a common location and dispersion parameter that correspond to the ensemble values of m and σ. To examine whether the ergodic specification AMIK CS yields an improvement in forecast accuracy, we compare the mean absolute forecast error of AMIK CS to AMIK TS and to the second best model, RW. Table 7 reports the incremental improvement in forecast accuracy relative to these two models as well as the p-value of the null hypothesis that the average loss difference between AMIK CS and, respectively, AMIK TS and RW, is equal to zero.[18]

[18] Although the structure of the models is different, our methodology is similar to that of Fairfield et al. (2009) who compare the forecasting performance of economy-wide and industry-specific dynamic panel models to that of the random walk. We will show later that our methodology clearly outperforms their dynamic panel models.

We observe that the ergodic model AMIK CS clearly outperforms the firm-specific model AMIK TS: the average difference in forecast accuracy is positive across forecast horizons, indicating that the ergodic specification leads to more accurate predictions than the firm-specific one. These differences are statistically significant at all conventional confidence levels. Gains relative to the random walk are even more pronounced and increase in size with the length of the time horizon, the only exception being the one-year ahead forecast for which the null hypothesis of equal predictive accuracy cannot be rejected. Hence, combining these results with those obtained in the previous subsection, we conclude that both the adjustment mechanism of the AMIK model and its ergodic parametrization lead to valuable gains in forecast accuracy compared to previously suggested models.

5.3 Comparison to alternative panel models

Finally, we check how the ergodic specification AMIK CS compares to panel models that have previously been employed in the profitability literature, and also aim to exploit cross-sectional information. To this end, we run a horse-race between AMIK CS, the structural partial adjustment model that we introduced in section 4.4, and the two non-linear partial adjustment models from section 4.5. As an additional robustness check, we also consider a cross-sectional modification of the Ornstein-Uhlenbeck process in which we estimate the location and dispersion parameter from pooled data assuming a stationary Gaussian distribution.

Average forecast errors and the SPA and MCS p-values are reported in Table 8. Except for the shortest time horizon, AMIK CS compares favorably to the competing panel models. Considering the size of the average forecast errors in Panel A, AMIK CS consistently outperforms OU CS, the structural partial adjustment model as well as the non-linear models with asymmetric adjustment for $h > 2$. Differences in forecast errors between the different models are relatively small, however, at least for the length of the available time series. This impression is also reflected in the SPA and MCS test results: AMIK CS, OU CS, and PAM pass the SPA test, yet we observe much lower p-values for the latter two; the SPA p-value of AMIK CS is equal to unity for $h > 2$, while it is considerably smaller for OU CS and PAM. Similarly, the model confidence set consists of AMIK CS, OU CS, and PAM. Again, AMIK CS exhibits the maximum probability that it belongs to the model confidence set among its competitors, while the p-values of the competing models are considerably smaller. Hence, we conclude that AMIK CS forecasts are more accurate than those of its competitors, but that the data are not informative enough to reject the null for all counterfactual models, especially at the short forecast horizon.

6 Discussion and conclusions

This paper shows that a parsimonious diffusion process forecasts corporate profitability significantly better than existing models. Not only does it outperform the random walk, which is no trivial task in small samples as we know since the seminal work of Meese and Rogoff (1983) (cf. Rossi (2006), as well as Rossi (2013) for a recent survey on exchange rate forecasting), it also outperforms partial adjustment models that use substantially more yet apparently superfluous information in their forecasts. The key elements in the predictive superiority of our model are the careful specification of the mean-reversion term in the diffusion process and the ergodicity of profit rates. The former is a direct consequence of the latter in the sense that the mean-reverting drift function of the diffusion process is constructed from the ensemble distribution of profit rates. While the different

Table 8: Forecast errors and statistical tests of relative forecast accuracy for panel models.

Horizon h	AMIK	OU	PAM	NLPAM EW	NLPAM IS
Panel A: Mean absolute forecast errors					
1	0.0301	0.0296	**0.0294**	0.0303	0.0298
2	0.0385	0.0386	**0.0382**	0.0403	0.0392
3	**0.0427**	0.0431	0.0431	0.0460	0.0442
4	**0.0450**	0.0454	0.0455	0.0490	0.0467
5	**0.0476**	0.0479	0.0484	0.0521	0.0494
6	**0.0492**	0.0494	0.0500	0.0544	0.0507
Panel B: P-values of the superior predictive ability test					
1	0.0102	**0.3098**	**1.0000**	0.0000	0.0220
2	**0.3372**	**0.3074**	**1.0000**	0.0000	0.0004
3	**1.0000**	**0.1366**	**0.3022**	0.0000	0.0128
4	**1.0000**	**0.1276**	**0.3230**	0.0002	0.0102
5	**1.0000**	**0.1016**	**0.2032**	0.0000	0.0166
6	**1.0000**	**0.2986**	**0.2490**	0.0000	0.0384
Panel C: P-values of the model confidence set					
1	0.0252	**0.6356**	**1.0000**	0.0000	0.0450
2	**0.6332**	**0.6332**	**1.0000**	0.0000	0.0348
3	**1.0000**	**0.4136**	**0.4136**	0.0000	0.0506
4	**1.0000**	**0.4476**	**0.4476**	0.0002	0.0510
5	**1.0000**	**0.3280**	**0.3280**	0.0000	0.0664
6	**1.0000**	**0.5620**	**0.5620**	0.0000	**0.1716**

Note: Panel A shows the mean absolute forecast errors (MAE) for pooled data. Bold values indicate the best model(s) in terms of the smallest average loss. In Panel B, we report the p-values of the test for superior predictive ability by Hansen (2005). The null hypothesis is that the respective model (i.e. each respective column) is not outperformed as a benchmark by any other of the four remaining models in terms of MAE. For each loss function p-values larger than 10 percent are shown in boldface. Panel C presents the MCS p-value for each competing model which captures the probability that this particular model belongs to the model confidence set introduced in Hansen et al. (2011). The null hypothesis is that all models that are part of the model confidence set have equal predictive accuracy in terms of MAE. In all three panels, AMIK and OU refer to the cross-sectional (CS) specification of the respective process where the estimates of the location and dispersion parameters are obtained from respectively fitting a Laplace or normal distribution to the data.

flavors of hitherto applied models account for the persistence of profit rates and their mean-reversion, they fail to account for their cross-sectional Laplace distribution and thus ignore valuable information that can evidently improve forecasting performance.

To realize these forecasting gains, however, one needs to part with several dearly held customs and intuitions that are prevalent in the time series and panel data literature. First notice that the diffusion process has a unit root (in its discretized version) yet it is stationary because of its drift function. Second, the specification of the drift function is essential and cannot be estimated, calibrated or correctly guessed from panel data models that ignore the distribution of profit rates. One can show that the dispersion of profit rates is determined by a ratio of the drift to the diffusion function (see Alfarano et al., 2012, for details), therefore the *noise* in the diffusion function is conceptually very different from an *error term* in conventional models because the noise captures systematically relevant fluctuations around the average rate of profit to which individual time series are reverting in the process of capital reallocation. Third, it seems that interest in unconditional scaling laws and distributional regularities more generally has considerably faded after the criticism by Brock (1999). Here we have translated the Laplace ensemble distribution of profit rates, that is an unconditional empirical regularity, into a conditional dynamic relationship that forecasts the profitability of firms better than existing models. In this sense, we confirm Brock's conjecture that distributional regularities can indeed be conveniently used to discriminate between potential stochastic processes in order to improve the conditional predictability of econometric models. Hence the improved predictive power of our model is a direct measure of our methodology's success in overcoming the Brock critique.

Our comparison of relative predictive accuracy shows that popular tests of forecasting performance such as SPA and MCS are rather conservative and hardly able to identify counterfactual processes in both real and simulated data if one deviates from the large time series setting. Such deviations necessarily occur in the many analyses of data that are available at annual, semiannual, or at best quarterly frequencies. To remedy this lack of power, we extend the design of the bootstrap from a strict time series to a panel setting that resamples from both the time and the cross-sectional sphere since ergodicity implies that the variability across firms is an appropriate source of additional statistical variation. This ergodic modification of the original SPA and MCS test designs corroborates the accuracy of our approach since we cannot reject the hypotheses that our model produces smaller forecast errors than alternative models that have previously been used to predict firm profitability.

The exploitation of ergodicity leads to better forecasts and more powerful bootstrapping procedures, yet it will elicit concerns of survivorship bias once we understand that ergodicity logically requires the survival of the studied entities. Since the notion of ergodicity has its historical roots in equilibrium statistical mechanics and its associated conservation laws, it is retrospectively perhaps not overly surprising that this logical prerequisite of survival has only been discussed very recently in realms outside of equilibrium statistical mechanics, initially by Peters and Gell-Mann and then by Taleb, and is explicitly lacking in the many definitions of ergodicity across different disciplines. While survival bias is certainly a legitimate concern in many areas of empirical work that seek to establish the effects of explanatory variables on (a set of) outcomes, we fear that it can at times also cloud our view on important macroscopic issues. Given the granular nature of surviving capital and its substantial impact on aggregate fluctuations, combined with the fact that most capital does not "disappear" but rather gets renamed after a change of ownership, we believe that our focus on the profitability of surviving corporations is none the less instructive. In a companion paper, for instance, we show that corporate idiosyncrasies have no impact on profitability once corporate capital survives for about two decades (Mundt et al., 2018). This is starkly at odds with an extensive body of literature that does not clearly distinguish between short- and long-lived firms and instead arrives at contradictory findings regarding the significance and directional impact of idiosyncratic factors on profitability. In this sense, a conscientious focus on survivors can help to clarify the origins of these contradictory findings. In spite of the remarkable dynamics of surviving capital, it goes without saying that abnormally profitable trading strategies will still be hard and risky to construct with our methodology, exactly because it is conditional on survival, which itself is not predictable with high levels of confidence. Still we have demonstrated that the exploitation of ergodicity leads to considerable forecasting gains, and its application to other data should accomplish the same.

A Descriptive statistics

Table 9: P-values of normality and kurtosis tests.

Year	AD	CVM	KUI	JB ALM	SW	AG
1980	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1981	0.0000	0.0007	0.0000	0.0000	0.0000	0.0000
1982	0.0239	0.0356	0.0045	0.0000	0.0000	0.0000
1983	0.0019	0.0071	0.0000	0.0000	0.0000	0.0000
1984	**0.0698**	**0.0800**	0.0110	0.0000	0.0000	0.0000
1985	0.0161	0.0303	0.0013	0.0000	0.0000	0.0000
1986	0.0006	0.0015	0.0000	0.0000	0.0000	0.0000
1987	0.0001	0.0004	0.0000	0.0000	0.0000	0.0000
1988	0.0020	0.0047	0.0000	0.0000	0.0000	0.0000
1989	0.0343	0.0328	0.0001	0.0000	0.0000	0.0000
1990	0.0081	0.0099	0.0000	0.0000	0.0000	0.0000
1991	0.0223	0.0272	0.0001	0.0000	0.0000	0.0000
1992	0.0183	0.0225	0.0001	0.0000	0.0000	0.0000
1993	0.0025	0.0034	0.0000	0.0000	0.0000	0.0000
1994	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
1995	0.0018	0.0047	0.0000	0.0000	0.0000	0.0000
1996	0.0050	0.0090	0.0000	0.0000	0.0000	0.0000
1997	0.0040	0.0095	0.0000	0.0000	0.0000	0.0000
1998	0.0056	0.0134	0.0000	0.0000	0.0000	0.0000
1999	0.0114	0.0226	0.0002	0.0000	0.0000	0.0000
2000	0.0012	0.0043	0.0000	0.0000	0.0000	0.0000
2001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
2002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2005	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000
2006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2008	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2010	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2013	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2014	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2016	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Note: Abbreviations refer to AD: Anderson-Darling test, CVM: Cramér-Von Mises test, KUI: Kuiper test, JB ALM: Jarque-Bera adjusted Lagrange multiplier test, SW: Shapiro-Wilk test, and AG: Anscombe-Glynn test. P-values greater than 5 percent are shown in boldface. Entries equal to 0.0000 imply p-values $< 5 \times 10^{-5}$.

B Simulation results

Table 10: Firm-by-firm SPA and MCS test results for simulated data.

Horizon h	SPA					MCS				
Panel A: T = 37										
	AMIK TS	OU	ARMA	ARIMA	RW	AMIK TS	OU	ARMA	ARIMA	RW
1	360	**378**	264	174	360	410	**421**	341	281	390
2	341	351	268	134	**352**	389	**394**	335	286	388
3	316	**323**	256	117	312	364	**377**	336	265	364
4	**313**	**313**	263	106	312	**357**	355	314	252	344
5	297	**313**	254	101	285	349	**351**	297	218	327
6	279	**299**	257	98	300	333	**338**	314	217	330
	AMIK CS	OU	ARMA	ARIMA	RW	AMIK CS	OU	ARMA	ARIMA	RW
1	**421**	323	235	160	317	**443**	390	318	260	358
2	**393**	316	249	117	296	**419**	366	322	268	350
3	**387**	277	241	101	237	**417**	335	303	238	319
4	**373**	262	241	89	220	**400**	322	286	209	276
5	**368**	260	222	81	198	**389**	299	269	178	256
6	**359**	235	215	79	200	**383**	293	269	174	268
Panel B: T = 100										
	AMIK TS	OU	ARMA	ARIMA	RW	AMIK TS	OU	ARMA	ARIMA	RW
1	**380**	330	73	14	**380**	**417**	389	141	108	407
2	**369**	328	118	28	360	**409**	387	189	166	390
3	**355**	346	155	67	325	**406**	388	237	210	370
4	**347**	337	201	91	313	**397**	390	238	211	356
5	334	**352**	204	108	306	387	**396**	243	210	337
6	332	**350**	206	111	288	387	**390**	256	218	335
	AMIK CS	OU	ARMA	ARIMA	RW	AMIK CS	OU	ARMA	ARIMA	RW
1	**456**	229	62	12	302	**461**	321	112	83	353
2	**450**	230	92	34	256	**456**	320	155	129	328
3	**447**	230	126	64	229	**456**	308	164	131	291
4	**449**	225	166	80	210	**459**	299	179	142	264
5	**448**	224	163	101	183	**454**	284	172	131	241
6	**439**	217	162	109	174	**452**	276	168	123	229

Note: Results refer to 465 simulations of the AMIK process with parameters $m = 0.095$, $\sigma = 0.058$, and idiosyncratic diffusion coefficients D_i that are estimated from the Worldscope data. Starting values are random draws from a Laplace distribution with parameters m and σ. The time series length is either $T = 37$ as for the real data or $T = 100$. The length of the in-sample period used for training the model is fixed at 20 observations in both scenarios. Hence, in the second experiment, all 63 additional data points increase the length of the evaluation period so that we can assess the effect of the sample size on the power of the test by comparing the two experiments. Entries show the number of cases for which the respective null hypothesis cannot be rejected on the 10 percent level. For each test the highest number of winnings is shown in boldface. OU refers to the time series specification of the Ornstein-Uhlenbeck process.

C Estimation results

Table 11: Summary statistics of the estimated parameters of AMIK TS.

Period	Min.	25%	Median	75%	Max.
			$\hat{m}$		
1980-1999	-0.0489	0.0737	0.1004	0.1367	0.3700
1981-2000	-0.0718	0.0716	0.0992	0.1359	0.3539
1982-2001	-0.0958	0.0703	0.0967	0.1334	0.3539
1983-2002	-0.1223	0.0695	0.0944	0.1313	0.3427
1984-2003	-0.1395	0.0700	0.0926	0.1299	0.3372
1985-2004	-0.1395	0.0695	0.0912	0.1264	0.3372
1986-2005	-0.1631	0.0679	0.0910	0.1259	0.3309
1987-2006	-0.1814	0.0691	0.0904	0.1234	0.3565
1988-2007	-0.2014	0.0679	0.0905	0.1238	0.3667
1989-2008	-0.2205	0.0670	0.0903	0.1214	0.3893
1990-2009	-0.2170	0.0653	0.0891	0.1184	0.3996
1991-2010	-0.1962	0.0648	0.0877	0.1176	0.4156
			$\hat{\sigma}$		
1980-1999	0.0035	0.0222	0.0370	0.0567	0.2963
1981-2000	0.0030	0.0218	0.0366	0.0556	0.2254
1982-2001	0.0040	0.0221	0.0366	0.0555	0.2363
1983-2002	0.0040	0.0222	0.0357	0.0533	0.1992
1984-2003	0.0045	0.0223	0.0348	0.0553	0.1850
1985-2004	0.0043	0.0222	0.0341	0.0547	0.2122
1986-2005	0.0037	0.0221	0.0345	0.0547	0.1884
1987-2006	0.0032	0.0218	0.0354	0.0543	0.3131
1988-2007	0.0028	0.0229	0.0351	0.0557	0.3918
1989-2008	0.0028	0.0232	0.0354	0.0573	0.3508
1990-2009	0.0027	0.0236	0.0375	0.0589	0.3047
1991-2010	0.0027	0.0232	0.0368	0.0585	0.3077
			$\hat{D}$		
1980-1999	0.0000	0.0005	0.0014	0.0040	27.6056
1981-2000	0.0000	0.0005	0.0014	0.0041	27.1916
1982-2001	0.0000	0.0005	0.0014	0.0039	28.6907
1983-2002	0.0000	0.0005	0.0014	0.0038	63.9495
1984-2003	0.0000	0.0005	0.0012	0.0032	50.2677
1985-2004	0.0000	0.0005	0.0012	0.0031	42.4566
1986-2005	0.0000	0.0005	0.0011	0.0029	24.9553
1987-2006	0.0000	0.0004	0.0010	0.0026	23.8767
1988-2007	0.0000	0.0004	0.0010	0.0027	77.9973
1989-2008	0.0000	0.0004	0.0010	0.0028	37.8642
1990-2009	0.0000	0.0004	0.0011	0.0029	35.1508
1991-2010	0.0000	0.0004	0.0011	0.0028	78.4933

Note: Entries equal to 0.0000 imply that the corresponding values are smaller than 5×10^{-5}.

Table 12: Summary statistics of the estimated parameters of AMIK CS.

Period	Min.	25%	Median	75%	Max.
			$\hat{D}$		
1980-1999	0.0000	0.0004	0.0010	0.0027	9.6918
1981-2000	0.0000	0.0004	0.0010	0.0026	6.9888
1982-2001	0.0000	0.0004	0.0011	0.0026	0.3425
1983-2002	0.0000	0.0004	0.0010	0.0025	0.1601
1984-2003	0.0000	0.0004	0.0009	0.0023	2.2816
1985-2004	0.0000	0.0004	0.0009	0.0023	1.5836
1986-2005	0.0000	0.0004	0.0009	0.0023	1.5836
1987-2006	0.0000	0.0004	0.0008	0.0021	1.8736
1988-2007	0.0000	0.0004	0.0008	0.0021	12.5901
1989-2008	0.0000	0.0003	0.0008	0.0022	9.6918
1990-2009	0.0000	0.0003	0.0009	0.0025	1.3057
1991-2010	0.0000	0.0003	0.0009	0.0023	9.6918

Note: The location parameter m and the dispersion parameter σ are fixed at their phenomenological values, summarized in Table 4 in the main text. Entries equal to 0.0000 imply that the corresponding estimates are smaller than 5×10^{-5}.

Table 13: Summary statistics of the estimated parameters of OU TS.

Period	Min.	25%	Median	75%	Max.
			$\hat{m}$		
1980-1999	-0.0514	0.0736	0.1037	0.1400	0.4210
1981-2000	-0.0649	0.0717	0.1010	0.1382	0.3799
1982-2001	-0.0345	0.0689	0.0974	0.1317	0.4038
1983-2002	-0.1391	0.0693	0.0968	0.1315	0.4265
1984-2003	-0.1586	0.0673	0.0957	0.1295	0.4017
1985-2004	-0.1475	0.0674	0.0940	0.1263	0.3954
1986-2005	-0.1935	0.0664	0.0930	0.1270	0.3727
1987-2006	-0.2091	0.0685	0.0932	0.1235	0.4156
1988-2007	-0.2710	0.0672	0.0917	0.1236	0.4088
1989-2008	-0.7252	0.0662	0.0906	0.1222	0.3997
1990-2009	-0.2813	0.0622	0.0873	0.1172	0.4076
1991-2010	-0.3333	0.0627	0.0865	0.1182	0.4295
			$\hat{\sigma}$		
1980-1999	0.0071	0.0261	0.0377	0.0599	0.4101
1981-2000	0.0062	0.0248	0.0381	0.0578	0.4267
1982-2001	0.0052	0.0246	0.0395	0.0574	0.2580
1983-2002	0.0046	0.0251	0.0373	0.0558	0.2989
1984-2003	0.0060	0.0249	0.0381	0.0576	0.2587
1985-2004	0.0060	0.0243	0.0369	0.0584	0.2374
1986-2005	0.0056	0.0247	0.0371	0.0583	0.2520
1987-2006	0.0047	0.0242	0.0365	0.0577	0.2854
1988-2007	0.0056	0.0242	0.0367	0.0576	0.3695
1989-2008	0.0040	0.0243	0.0372	0.0580	0.9164
1990-2009	0.0048	0.0239	0.0370	0.0583	0.5244
1991-2010	0.0038	0.0233	0.0374	0.0566	0.5416
			$\hat{D}$		
1980-1999	0.0000	0.0005	0.0014	0.0037	0.4382
1981-2000	0.0000	0.0005	0.0014	0.0038	6.9470
1982-2001	0.0000	0.0005	0.0013	0.0036	3.6951
1983-2002	0.0000	0.0005	0.0014	0.0031	0.4925
1984-2003	0.0000	0.0005	0.0012	0.0029	0.1651
1985-2004	0.0000	0.0005	0.0011	0.0027	3.0114
1986-2005	0.0000	0.0005	0.0011	0.0028	0.3227
1987-2006	0.0000	0.0004	0.0011	0.0026	1.6097
1988-2007	0.0000	0.0004	0.0010	0.0025	0.3960
1989-2008	0.0000	0.0004	0.0010	0.0027	0.9644
1990-2009	0.0000	0.0004	0.0011	0.0029	0.3035
1991-2010	0.0000	0.0004	0.0011	0.0028	4.1833

Note: Entries equal to 0.0000 imply that the corresponding values are smaller than 5×10^{-5}.

Table 14: Summary statistics of the estimated parameters of OU CS.

Period	Min.	25%	Median	75%	Max.
			$\hat{D}$		
1980-1999	0.0000	0.0004	0.0010	0.0028	0.0250
1981-2000	0.0000	0.0004	0.0010	0.0028	0.0225
1982-2001	0.0000	0.0004	0.0010	0.0025	0.0273
1983-2002	0.0000	0.0004	0.0010	0.0024	0.0206
1984-2003	0.0000	0.0004	0.0009	0.0023	0.0349
1985-2004	0.0000	0.0004	0.0009	0.0023	0.0207
1986-2005	0.0000	0.0004	0.0009	0.0022	0.0207
1987-2006	0.0000	0.0004	0.0008	0.0021	0.0261
1988-2007	0.0000	0.0003	0.0008	0.0020	0.0279
1989-2008	0.0000	0.0003	0.0008	0.0021	0.0320
1990-2009	0.0000	0.0003	0.0009	0.0024	0.0355
1991-2010	0.0000	0.0003	0.0008	0.0023	0.0479

Note: The location parameter m and the dispersion parameter σ are estimated from Gaussian distributions that are fitted to pooled data. Entries equal to 0.0000 imply that the corresponding estimates are smaller than 5×10^{-5}.

Table 15: Summary statistics for the estimated parameters of ARMA.

Period	Min.	25%	Median	75%	Max.
			$\hat{c}$		
1980-1999	-0.9227	0.0272	0.0515	0.0845	0.9339
1981-2000	-104.7131	0.0265	0.0481	0.0827	4.6861
1982-2001	-10.4311	0.0275	0.0481	0.0821	1.0418
1983-2002	-122.2397	0.0267	0.0484	0.0834	0.5198
1984-2003	-1.2884	0.0259	0.0473	0.0802	0.9154
1985-2004	-2.9854	0.0261	0.0455	0.0754	7.9124
1986-2005	-2.6578	0.0256	0.0426	0.0731	2.5900
1987-2006	-0.5794	0.0237	0.0431	0.0697	0.8173
1988-2007	-1.1162	0.0232	0.0418	0.0676	9.0449
1989-2008	-7.0722	0.0224	0.0394	0.0672	1.8875
1990-2009	-1.9883	0.0220	0.0403	0.0664	11.1863
1991-2010	-2.3408	0.0212	0.0404	0.0671	9.9020
			$\hat{\lambda}$		
1980-1999	-42.8141	0.1498	0.5025	0.6903	10.9541
1981-2000	-76.4585	0.1493	0.5061	0.7058	1182.9813
1982-2001	-7.5930	0.0999	0.4974	0.6975	156.6820
1983-2002	-425.9735	0.1319	0.4942	0.6721	2102.9842
1984-2003	-7.8685	0.1605	0.4893	0.7042	15.0287
1985-2004	-75.0235	0.1978	0.5124	0.6854	46.9542
1986-2005	-33.0459	0.2100	0.5219	0.6888	18.0523
1987-2006	-8.5463	0.2636	0.5350	0.7097	8.2472
1988-2007	-96.8263	0.2857	0.5296	0.7100	12.7770
1989-2008	-18.6911	0.2915	0.5438	0.7165	108.8221
1990-2009	-148.6684	0.2922	0.5510	0.7193	22.6550
1991-2010	-216.5744	0.2827	0.5481	0.7182	42.7274
			$\hat{\psi}$		
1980-1999	-14.9776	-0.0538	0.2002	0.5437	42.8123
1981-2000	-1182.9816	-0.0676	0.1968	0.5515	76.4629
1982-2001	-156.6805	-0.0626	0.2140	0.5806	7.6365
1983-2002	-2102.9835	-0.0107	0.2584	0.5556	425.9736
1984-2003	-15.0498	-0.0419	0.1928	0.5137	7.9088
1985-2004	-46.9609	-0.0229	0.1945	0.4754	75.0284
1986-2005	-18.0662	-0.0404	0.2149	0.5069	33.0370
1987-2006	-8.2253	-0.0300	0.2135	0.5087	8.5292
1988-2007	-12.8117	-0.0258	0.2049	0.4988	96.8289
1989-2008	-108.8210	-0.0226	0.1703	0.4757	18.7106
1990-2009	-22.6645	-0.0697	0.1834	0.4915	148.6667
1991-2010	-42.7190	-0.0446	0.1903	0.4979	216.5742
			$\hat{\sigma}^2$		
1980-1999	0.0000	0.0004	0.0008	0.0019	0.0843
1981-2000	0.0000	0.0004	0.0008	0.0020	0.0769
1982-2001	0.0000	0.0004	0.0009	0.0019	0.0426
1983-2002	0.0000	0.0004	0.0008	0.0017	0.0454
1984-2003	0.0000	0.0004	0.0008	0.0017	0.0353
1985-2004	0.0000	0.0003	0.0008	0.0017	0.0352
1986-2005	0.0000	0.0003	0.0008	0.0017	0.0303
1987-2006	0.0000	0.0003	0.0007	0.0016	0.0348
1988-2007	0.0000	0.0003	0.0007	0.0016	0.0297
1989-2008	0.0000	0.0003	0.0007	0.0016	0.2208
1990-2009	0.0000	0.0003	0.0008	0.0018	0.1902
1991-2010	0.0000	0.0003	0.0008	0.0016	0.1796

Note: Entries equal to 0.0000 imply that the corresponding values are smaller than 5×10^{-5}.

Table 16: Summary statistics for the estimated parameters of ARIMA.

Period	Min.	25%	Median	75%	Max.
			$\hat{\lambda}$		
1980-1999	-211.0938	-0.5234	0.1014	0.7645	174.8884
1981-2000	-80.4232	-0.5979	0.1289	0.7736	22.4037
1982-2001	-160.4340	-0.5599	0.1647	0.7962	1089.6294
1983-2002	-353.2678	-0.4222	0.2198	0.8769	110.0544
1984-2003	-47154.0926	-0.5238	0.1389	0.7083	1054.2162
1985-2004	-108.2522	-0.5494	0.1340	0.7729	51.5886
1986-2005	-68.8545	-0.5947	0.0721	0.7039	46.1794
1987-2006	-71.2947	-0.5500	0.0640	0.6925	174.8354
1988-2007	-68.6584	-0.5735	0.0575	0.7256	149.5186
1989-2008	-2892.6316	-0.5915	0.0843	0.7117	612.8428
1990-2009	-99.5175	-0.5716	0.1753	0.7330	223.8485
1991-2010	-51.6857	-0.4959	0.1667	0.6779	419.4589
			$\hat{\psi}$		
1980-1999	-174.8887	-1.0020	-0.4312	0.5400	211.0947
1981-2000	-22.3862	-1.0006	-0.4645	0.6494	80.4274
1982-2001	-1089.6297	-1.0011	-0.5632	0.4500	160.4363
1983-2002	-110.0584	-1.1239	-0.6171	0.2780	353.2683
1984-2003	-1054.2164	-1.0000	-0.4398	0.4481	47154.0926
1985-2004	-51.5841	-1.0120	-0.5480	0.4760	108.2533
1986-2005	-46.1846	-1.0003	-0.3937	0.6226	68.8601
1987-2006	-174.8356	-1.0003	-0.3637	0.6152	71.3003
1988-2007	-149.5205	-1.0000	-0.3103	0.6363	68.6620
1989-2008	-612.8425	-1.0000	-0.3210	0.6858	2892.6317
1990-2009	-223.8490	-1.0000	-0.4303	0.5852	99.5186
1991-2010	-419.4585	-1.0000	-0.4543	0.4831	51.6785
			$\hat{\sigma}^2$		
1980-1999	0.0000	0.0004	0.0010	0.0021	0.0831
1981-2000	0.0000	0.0004	0.0009	0.0021	0.0570
1982-2001	0.0000	0.0004	0.0010	0.0020	0.0398
1983-2002	0.0000	0.0004	0.0009	0.0019	0.0383
1984-2003	0.0000	0.0004	0.0009	0.0018	0.0360
1985-2004	0.0000	0.0003	0.0008	0.0017	0.0523
1986-2005	0.0000	0.0003	0.0008	0.0017	0.0442
1987-2006	0.0000	0.0003	0.0008	0.0016	0.0440
1988-2007	0.0000	0.0003	0.0008	0.0017	0.0420
1989-2008	0.0000	0.0003	0.0008	0.0017	0.1113
1990-2009	0.0000	0.0003	0.0008	0.0018	0.1735
1991-2010	0.0000	0.0003	0.0008	0.0018	0.1686

Note: Entries equal to 0.0000 imply that the corresponding values are smaller than 5×10^{-5}.

Table 17: Estimation results for the structural partial adjustment model.

Period	Parameter								Diagnostics		
	$\hat{\delta}$		$\hat{\zeta}$		$\hat{\eta}$		$\hat{\theta}$		Adj. R^2	AIC	BIC
	Mean	t-stat	Mean	t-stat	Mean	t-stat	Mean	t-stat			
1980-1999	0.0570	12.2559	0.0485	13.9470	-0.0458	-12.1026	0.0509	2.0521	0.31	-1160	-1139
1981-2000	0.0575	12.0871	0.0453	14.2058	-0.0450	-12.2554	0.0532	2.1672	0.30	-1167	-1147
1982-2001	0.0543	16.0287	0.0456	14.7973	-0.0440	-12.5042	0.0581	2.4624	0.31	-1175	-1154
1983-2002	0.0521	14.6749	0.0451	15.1058	-0.0428	-13.0826	0.0577	2.4420	0.32	-1185	-1164
1984-2003	0.0507	13.514	0.0456	14.5828	-0.0420	-13.5446	0.0440	1.8508	0.34	-1200	-1179
1985-2004	0.0493	13.2696	0.0450	14.8110	-0.0414	-13.8070	0.0435	1.8340	0.34	-1205	-1185
1986-2005	0.0489	13.2323	0.0449	14.7818	-0.0422	-13.9676	0.0296	1.3828	0.35	-1211	-1191
1987-2006	0.0481	13.2291	0.0463	16.8377	-0.0431	-13.9114	0.0091	1.1505	0.36	-1224	-1203
1988-2007	0.0468	13.6708	0.0467	17.4995	-0.0418	-16.0550	0.0098	1.2247	0.36	-1223	-1202
1989-2008	0.0470	13.5691	0.0454	16.3277	-0.0432	-12.8220	0.0113	1.3713	0.35	-1194	-1173
1990-2009	0.0446	12.7046	0.0470	15.2928	-0.0429	-12.8394	0.0113	1.3725	0.36	-1193	-1172
1991-2010	0.0423	12.4935	0.0487	15.5559	-0.0423	-12.7385	0.0103	1.2663	0.37	-1189	-1168
1980-1999	0.0545	10.8677	0.0479	14.2491					0.27	-1137	-1125
1981-2000	0.0550	10.7871	0.0448	14.1500					0.26	-1143	-1131
1982-2001	0.0517	12.8823	0.0454	15.2696					0.27	-1150	-1138
1983-2002	0.0488	11.4514	0.0454	15.2635					0.28	-1160	-1148
1984-2003	0.0459	10.0983	0.0465	14.5443					0.30	-1175	-1163
1985-2004	0.0439	9.8764	0.0463	14.5850					0.30	-1180	-1168
1986-2005	0.0425	9.8998	0.0463	14.5788					0.30	-1186	-1174
1987-2006	0.0403	10.4836	0.0477	16.6568					0.32	-1198	-1186
1988-2007	0.0390	10.6520	0.0481	17.3283					0.32	-1198	-1186
1989-2008	0.0388	10.6538	0.0469	16.4462					0.31	-1170	-1157
1990-2009	0.0361	9.5563	0.0487	15.0949					0.32	-1169	-1157
1991-2010	0.0336	9.2062	0.0504	15.2350					0.33	-1166	-1154

Period	$\hat{\alpha}$		$\hat{\beta}$		$\hat{\gamma}$		Adj. R^2	AIC	BIC
	Mean	t-stat	Mean	t-stat	Mean	t-stat			
1980-1999	-0.0020	-1.2752	-0.1935	-10.5650	-0.0880	-3.2571	0.12	-1499	-1482
1981-2000	-0.0010	-0.6554	-0.1990	-10.7266	-0.0767	-2.6406	0.12	-1506	-1490
1982-2001	-0.0018	-1.0967	-0.1895	-9.8582	-0.0947	-3.3168	0.12	-1512	-1495
1983-2002	-0.0025	-1.8674	-0.1804	-9.6071	-0.0869	-2.9806	0.11	-1513	-1497
1984-2003	-0.0022	-1.6064	-0.1765	-9.3800	-0.0892	-3.0615	0.11	-1539	-1523
1985-2004	-0.0008	-0.6077	-0.1931	-8.5572	-0.0688	-1.8502	0.12	-1548	-1531
1986-2005	-0.0002	-0.1395	-0.1929	-8.4344	-0.0662	-1.7988	0.12	-1560	-1544
1987-2006	0.0000	-0.0197	-0.1741	-7.1469	-0.0650	-1.7617	0.11	-1590	-1574
1988-2007	-0.0007	-0.4611	-0.1636	-6.7560	-0.0494	-1.2655	0.10	-1588	-1571
1989-2008	-0.0004	-0.2924	-0.1389	-3.4327	-0.0554	-1.3888	0.10	-1553	-1537
1990-2009	-0.0013	-0.6435	-0.1541	-3.4895	-0.0558	-1.3996	0.13	-1539	-1522
1991-2010	-0.0006	-0.2850	-0.1630	-3.5914	-0.0629	-1.5471	0.14	-1521	-1505
1980-1999	-0.0020	-1.4872	-0.2228	-13.3675			0.11	-1496	-1484
1981-2000	-0.0010	-0.7668	-0.2270	-13.6461			0.11	-1504	-1492
1982-2001	-0.0016	-1.0716	-0.2220	-12.7891			0.11	-1509	-1496
1983-2002	-0.0024	-1.8441	-0.2109	-12.0789			0.10	-1510	-1498
1984-2003	-0.0020	-1.5319	-0.2067	-11.7716			0.10	-1535	-1523
1985-2004	-0.0005	-0.4249	-0.2220	-11.7500			0.11	-1543	-1531
1986-2005	-0.0002	-0.1878	-0.2210	-11.6781			0.10	-1555	-1543
1987-2006	-0.0001	-0.0833	-0.2012	-9.7496			0.09	-1584	-1572
1988-2007	-0.0007	-0.5397	-0.1859	-8.8604			0.08	-1581	-1569
1989-2008	-0.0005	-0.3513	-0.1568	-4.3426			0.09	-1546	-1533
1990-2009	-0.0013	-0.7377	-0.1739	-4.2595			0.11	-1529	-1517
1991-2010	-0.0005	-0.2440	-0.1869	-4.3538			0.12	-1513	-1501

Note: Means and t-statistics for the means of the year-by-year cross-section regression coefficients. Intercepts and slopes are averaged across 20 years of each in-sample period. The t-statistics of these averages are computed as the ratio of the mean coefficient to its time series standard deviation multiplied with $\sqrt{20}$ (Fama and French, 2000).

Table 18: Estimation results for the non-linear partial adjustment models.

Period / Sector	Parameter					
	$\hat{\alpha}$		$\hat{\beta}$		$\hat{\gamma}$	
	Mean	t-stat	Mean	t-stat	Mean	t-stat
Panel A: Economy-wide model NLPAM EW						
1980-1999	0.0249	7.3087	0.7926	39.3361	-0.1230	-5.5888
1981-2000	0.0231	7.5537	0.8036	43.6785	-0.1201	-5.5831
1982-2001	0.0223	7.2567	0.8131	46.7104	-0.1158	-5.2789
1983-2002	0.0213	6.7725	0.8120	46.5730	-0.1171	-5.3451
1984-2003	0.0186	6.6330	0.8229	51.7984	-0.1019	-3.9537
1985-2004	0.0186	6.5903	0.8266	51.2232	-0.1026	-3.9845
1986-2005	0.0206	7.8254	0.8213	52.2331	-0.1273	-4.6207
1987-2006	0.0198	7.7694	0.8259	54.1198	-0.1189	-4.3058
1988-2007	0.0184	8.5848	0.8387	66.7031	-0.1145	-4.2545
1989-2008	0.0163	7.4765	0.8491	70.6096	-0.0939	-3.3200
1990-2009	0.0141	4.3936	0.8623	48.2717	-0.0587	-1.2660
1991-2010	0.0142	4.4264	0.8547	42.5397	-0.0665	-1.3962
Panel B: Industry-specific model NLPAM IS						
Mining	0.0119	1.5807	0.8173	10.2598	0.0588	0.5843
Construction	0.0203	1.6033	0.6823	5.9691	0.1660	0.6230
Manufacturing	0.0221	5.6347	0.8120	38.4352	-0.1009	-2.6791
Transportation and public utilities	0.0240	5.6151	0.6999	14.9160	-0.0253	-1.5130
Wholesale trade	0.0127	2.0044	0.8786	24.2021	0.0151	0.2574
Retail trade	0.0097	1.5169	0.9064	22.7879	-0.0015	0.0684
Finance, insurance, and real estate	0.0037	1.2818	0.9845	16.2656	-0.1220	-1.6807
Services	0.0136	3.9156	0.8922	27.8974	-0.1202	-3.5674

Note: Means and t-statistics for the means of the year-by-year cross-section regression coefficients. Intercepts and slopes are averaged across the 20 years of each in-sample period. The t-statistics of these averages are computed as the ratio of the mean coefficient to its time series standard deviation multiplied with $\sqrt{20}$.

References

Albrecht, S., Lookabill, L., and McKeown, J. (1977). The time series properties of annual earnings. *Journal of Accounting Research*, 15(2):226–244.

Alfarano, S. and Milaković, M. (2008). Does classical competition explain the statistical features of firm growth? *Economics Letters*, 101(3):272–274.

Alfarano, S., Milaković, M., Irle, A., and Kauschke, J. (2012). A statistical equilibrium model of competitive firms. *Journal of Economic Dynamics and Control*, 36(1):136–149.

Allen, D. E. and Salim, H. M. (2005). Forecasting profitability and earnings: a study of the uk market (1982–2000). *Applied Economics*, 37(17):2009–2018.

Anscombe, F. and Glynn, W. (1983). Distribution of kurtosis statistic for normal statistics. *Biometrika*, 70(1):227–234.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science*, 293(5536):1818–1820.

Ball, R. and Watts, R. (1972). Some time series properties of accounting income. *The Journal of Finance*, 27(3):663–681.

Bauman, M. (2014). Forecasting operating profitability with DuPont analysis: further evidence. *Review of Accounting and Finance*, 13(2):191–205.

Bottazzi, G. and Secchi, A. (2003a). Common properties and sectoral specificities in the dynamics of U.S. manufacturing companies. *Review of Industrial Organization*, 23(3):217–232.

Bottazzi, G. and Secchi, A. (2003b). Why are distributions of firm growth rates tent-shaped? *Economics Letters*, 80(3):415–420.

Bottazzi, G. and Secchi, A. (2006). Explaining the distribution of firm growth rates. *The RAND Journal of Economics*, 37(2):235–256.

Box, G. and Pierce, D. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332):1509–1526.

Brock, W. A. (1999). Scaling in economics: a reader's guide. *Industrial and Corporate Change*, 8(3):409–446.

Brown, L. and Rozeff, M. (1979). Univariate time-series models of quarterly accounting earnings per share: a proposed model. *Journal of Accounting Research*, 17(1):179–189.

Burgstahler, D. and Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, 24(1):99–126.

Callen, J., Cheung, S., Kwan, C., and Yip, P. (1993). An empirical investigation of the random character of annual earnings. *Journal of Accounting, Auditing & Finance*, 8(2):151–162.

Cochrane, J. (1991). A critique of the application of unit root tests. *Journal of Economic Dynamics and Control*, 15(2):275–284.

Collins, W. and Hopwood, W. (1980). A multivariate analysis of annual earnings forecasts generated from quarterly forecasts of financial analysts and univariate time series models. *Journal of Accounting Research*, 18(2):390–406.

Corradi, V. and Distaso, W. (2011). Multiple forecast model evaluation. In Clements, M. and Hendry, D., editors, *The Oxford handbook of economic forecasting*, pages 391–413. Oxford University Press, New York, NY.

Cox, J., Ingersoll, J., and Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.

Daepp, M. I. G., Hamilton, M. J., West, G. B., and Bettencourt, L. M. A. (2015). The mortality of companies. *Journal of The Royal Society Interface*, 12(106).

Diebold, F. (2006). *Elements of forecasting*. South Western, Mason, Ohio.

Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

Erlingsson, E. J., Alfarano, S., Raberto, M., and Stefánsson, H. (2012). On the distributional properties of size, profit and growth of icelandic firms. *Journal of Economic Interaction and Coordination*, April.

Evans, M., Njoroge, K., and Yong, K. (2017). An examination of the statistical significance and economic relevance of profitability and earnings forecasts from models and analysts. *Contemporary Accounting Research*, 34(3):1453–1488.

Fagiolo, G., Alessi, L., Barigozzi, M., and Capasso, M. (2010). On the distributional properties of household consumption expenditures: the case of Italy. *Empirical Economics*, 38(3):717–741.

Fagiolo, G., Napoletano, M., and Roventini, A. (2008). Are output growth-rate distributions fat-tailed? Some evidence from OECD countries. *Journal of Applied Econometrics*, 23(5):639–669.

Fairfield, P., Ramnath, S., and Yohn, T. (2009). Do industry-level analyses improve forecasts of financial performance? *Journal of Accounting Research*, 47(1):147–176.

Fairfield, P., Sweeney, R., and Yohn, T. (1996). Accounting classification and the predictive content of earnings. *The Accounting Review*, 71(3):337–355s.

Fairfield, P. and Yohn, T. (2001). Using asset turnover and profit margin to forecast changes in profitability. *Review of Accounting Studies*, 6(4):371–385.

Fama, E. and French, K. (2001). Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial Economics*, 60(1):3–43.

Fama, E. F. and French, K. (2000). Forecasting profitability and earnings. *Journal of Business*, 73(2):161–178.

Fama, E. F. and French, K. (2002). Testing trade-off and pecking order predictions about dividends and debt. *Review of Financial Studies*, 15(1):1–33.

Feltham, G. and Ohlson, J. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research*, 11(2):689–731.

Foster, G. (1977). Quartely accounting data: time-series properties and predictive-ability results. *The Accounting Review*, 52(1):1–21.

Freeman, R., Ohlson, J., and Penman, S. (1982). Book rate-of-return and prediction of earnings changes: an empirical investigation. *Journal of Accounting Research*, 20(2):639–653.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.

Gallizo, J. L., Gargallo, P., and Salvador, M. (2008). Multivariate partial adjustment of financial ratios: a Bayesian hierarchical approach. *Journal of Applied Econometrics*, 23(1):43–64.

Gardiner, C. (2009). *Stochastic methods - A handbook for the natural and social sciences*. Springer, Berlin.

Garibaldi, U. and Scalas, E. (2010). *Finitary probabilistic methods in econophysics*. Cambridge University Press, Cambridge, UK.

Geroski, P., Lazarova, S., Urga, G., and Walters, C. (2003). Are differences in firm size transitory or permanent? *Journal of Applied Econometrics*, 18(1):47–59.

Geroski, P. A. and Jacquemin, A. (1988). The persistence of profits: a European comparison. *The Economic Journal*, 98(391):375–389.

Ghongadze, J. and Lux, T. (2012). Modelling the dynamics of EU economic sentiment indicators: an interaction-based approach. *Applied Economics*, 44(24):3065–3088.

Giacomini, R. and Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, 25(4):595–620.

Goddard, J. and Wilson, J. O. S. (1999). The persistence of profit: A new empirical interpretation. *International Journal of Industrial Organization*, 17(5):663–687.

Griffin, P. (1977). The time-series behavior of quarterly earnings: preliminary evidence. *Journal of Accounting Research*, 15(1):71–83.

Gschwandtner, A. (2005). Profit persistence in the very long run: evidence from survivors and exciters. *Applied Economics*, 37(7):793–806.

Hansen, P. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23(4):365–380.

Hansen, P., Lunde, A., and Nason, J. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Hou, K., van Dijk, M., and Zhang, Y. (2012). The implied cost of capital: a new approach. *Journal of Accounting and Economics*, 53(3):504–526.

Kapetanios, G. (2008). A bootstrap procedure for panel data sets with many cross-sectional units. *Econometrics Journal*, 11(2):377–395.

Kotz, S., Kozubowski, T., and Podgórski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering, and finance*. Birkhäuser, New York, NY.

Lintner, J. and Glauber, R. (1978). Higgledy-piggledy growth in America. In Lorie, J., editor, *Modern developments in investment management*, pages 594–611. Thomson Learning, Hinsdale, IL.

Little, I. (1962). Higgledy-piggledy growth. *Bulletin of the Oxford University Institute of Economics and Statistics*, 24(4):387–412.

Ljung, G. and Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Lookabill, L. (1976). Some additional evidence on the time series properties of accounting earnings. *The Accounting Review*, 51(4):724–738.

Lux, T. (2009). Rational forecasts or social opinion dynamics? identification of interaction effects in a business climate survey. *Journal of Economic Behavior and Organization*, 72(2):638–655.

Machado, J. and Mata, J. (2000). Box-Cox quantile regression and the distribution of firm sizes. *Journal of Applied Econometrics*, 15(3):253–274.

Meese, R. A. and Rogoff, K. (1983). Empirical exchange rate models of the Seventies: do they fit out of sample? *Journal of International Economics*, 14:3–24.

Mittelhammer, R. (2013). *Mathematical statistics for economics and business*. Springer, New York, NY.

Mueller, D. C. (1977). The persistence of profits above the norm. *Economica*, 44(176):369–380.

Mueller, D. C. (1986). *Profits in the long run*. Cambridge University Press, Cambridge, MA.

Mueller, D. C. (1990). The persistence of profits in the United States. In Mueller, D. C., editor, *The dynamics of company profits*, pages 35–58. Cambridge University Press.

Mundt, P., Alfarano, S., and Milaković, M. (2016). Gibrat's law redux: think profitability instead of growth. *Industrial and Corporate Change*, 25(4):549–571.

Mundt, P., Milaković, M., and Alfarano, S. (2018). The ultimate corporate objective is survival. Unpublished manuscript.

Nissim, D. and Penman, S. (2001). Ratio analysis and equity valuation: from research to practice. *Review of Accounting Studies*, 6(1):109–154.

Ohlson, J. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research*, 11(2):661–687.

Peters, O. and Gell-Mann, M. (2016). Evaluating gambles using dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(2):023103.

Politis, D. and Romano, J. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

Priestley, M. and Rao, T. S. (1969). A test for non-stationarity of time series. *Journal of the Royal Statistical Society*, 31(1):140–149.

Reed, W. (2001). The Pareto, Zipf and other power laws. *Economics Letters*, 74(1):15–19.

Reed, W. (2003). The Pareto law of incomes: an explanation and an extension. *Physica A*, 319(1):469–486.

Richardson, S., Tuna, I., and Wysocki, P. (2010). Accounting anomalies and fundamental analysis: a review of recent research advances. *Journal of Accounting and Economics*, 50(2-3):410–454.

Risken, H. (1996). *The Fokker-Planck equation: Methods of solutions and applications*. Springer, Berlin.

Rossi, B. (2006). Are exchange rates really random walks? Some evidence robust to parameter instability. *Macroeconomic Dynamics*, 10(1):20–38.

Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, 51(4):1063–1119.

Scharfenaker, E. and Semieniuk, G. (2016). A statistical equilibrium approach to the distribution of profit rates. *Metroeconomica*, in press.

Soliman, M. (2008). The use of DuPont analysis by market participants. *The Accounting Review*, 83(3):823–853.

Stanley, M. H. R., Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A., and Stanley, H. E. (1996). Scaling behaviour in the growth of companies. *Nature*, 379(29):804–806.

Stanley, M. H. R., Buldyrev, S. V., Havlin, S., Mantegna, R. N., Salinger, M. A., and Stanley, H. E. (1995). Zipf plots and the size distribution of firms. *Economics Letters*, 49(4):453–457.

Stigler, G. (1963). *Capital and rates of return in manufacturing industries*. Princeton University Press, Princeton, NJ.

Taleb, N. N. (2018). *Skin in the Game: Hidden Asymmetries in Daily Life*. Random House.

Toda, A. A. (2012). The double power law in income distribution: explanations and evidence. *Journal of Economic Behavior and Organization*, 84(1):364–381.

Toda, A. A. (2017). A note on the size distribution of consumption: more double Pareto than lognormal. *Macroeconomic Dynamics*, 21(6):1508–1518.

Toda, A. A. and Walsh, K. (2015). The double power law in consumption and implications for testing Euler equations. *Journal of Political Economy*, 123(5):1177–1200.

Toda, A. A. and Walsh, K. (2017). Fat tails and spurious estimation of consumption-based asset pricing models. *Journal of Applied Econometrics*, 32(6):1156–1177.

Uhlenbeck, G. and Ornstein, L. (1930). On the theory of brownian motion. *Physical Review*, 36(5):823–841.

Waring, G. F. (1996). Industry differences in the persistence of firm-specific returns. *American Economic Review*, 86(5):1253–65.

Watts, R. and Leftwich, R. (1977). The time series of annual accounting earnings. *Journal of Accounting Research*, 15(2):253–271.

White, H. (2000). A reality check for data snooping. *Econometrica*, 68(5):1097–1126.

BERG Working Paper Series (most recent publications)

113 Julio **González-Díaz**, Florian **Herold** and Diego **Domínguez**, Strategic Sequential Voting, July 2016

114 Stefanie Yvonne **Schmitt**, Rational Allocation of Attention in Decision-Making, July 2016

115 Florian **Herold** and Christoph **Kuzmics**, The evolution of taking roles, September 2016.

116 Lisa **Planer-Friedrich** and Marco **Sahm**, Why Firms Should Care for All Consumers, September 2016.

117 Christoph **March** and Marco **Sahm**, Asymmetric Discouragement in Asymmetric Contests, September 2016.

118 Marco **Sahm**, Advance-Purchase Financing of Projects with Few Buyers, October 2016.

119 Noemi **Schmitt** and Frank **Westerhoff**, On the bimodality of the distribution of the S&P 500's distortion: empirical evidence and theoretical explanations, January 2017

120 Marco **Sahm**, Risk Aversion and Prudence in Contests, March 2017

121 Marco **Sahm**, Are Sequential Round-Robin Tournaments Discriminatory?, March 2017

122 Noemi **Schmitt**, Jan **Tuinstra** and Frank **Westerhoff**, Stability and welfare effects of profit taxes within an evolutionary market interaction model, May 2017

123 Johanna Sophie **Quis** and Simon **Reif**, Health Effects of Instruction Intensity – Evidence from a Natural Experiment in German High-Schools, May 2017

124 Lisa **Planer-Friedrich** and Marco **Sahm**, Strategic Corporate Social Responsibility, May 2017

125 Peter **Flaschel**, Matthieu **Charpe**, Giorgos **Galanis**, Christian R. **Proaño** and Roberto **Veneziani**, Macroeconomic and Stock Market Interactions with Endogenous Aggregate Sentiment Dynamics, May 2017

126 Christian **Menden** and Christian R. **Proaño**, Dissecting the Financial Cycle with Dynamic Factor Models, May 2017

127 Christoph **March** and Marco **Sahm**, Contests as Selection Mechanisms: The Impact of Risk Aversion, July 2017

128 Ivonne **Blaurock**, Noemi **Schmitt** and Frank **Westerhoff**, Market entry waves and volatility outbursts in stock markets, August 2017

129 Christoph **Laica**, Arne **Lauber** and Marco **Sahm**, Sequential Round-Robin Tournaments with Multiple Prizes, September 2017

130 Joep **Lustenhouwer** and Kostas **Mavromatis**, Fiscal Consolidations and Finite Planning Horizons, December 2017

131 Cars **Hommes** and Joep **Lustenhouwer**, Managing Unanchored, Heterogeneous Expectations and Liquidity Traps, December 2017

132 Cars **Hommes**, Joep **Lustenhouwer** and Kostas **Mavromatis**, Fiscal Consolidations and Heterogeneous Expectations, December 2017

133 Roberto **Dieci**, Noemi **Schmitt** and Frank **Westerhoff**, Interactions between stock, bond and housing markets, January 2018

134 Noemi **Schmitt**, Heterogeneous expectations and asset price dynamics, January 2018

135 Carolin **Martin** and Frank **Westerhoff**, Regulating speculative housing markets via public housing construction programs: Insights from a heterogeneous agent model, May 2018

136 Roberto **Dieci**, Noemi **Schmitt** and Frank **Westerhoff**, Steady states, stability and bifurcations in multi-asset market models, July 2018

137 Steffen **Ahrens**, Joep **Lustenhouwer** and Michele **Tettamanzi**, The Stabilizing Role of Forward Guidance: A Macro Experiment, September 2018

138 Joep **Lustenhouwer**, Fiscal Stimulus in an Expectation Driven Liquidity Trap, September 2018

139 Tim **Hagenhoff**, An aggregate welfare optimizing interest rate rule under heterogeneous expectations, October 2018

140 Johanna Sophie **Quis**, Anika **Bela** and Guido **Heineck**, Preschoolers' self-regulation, skill differentials, and early educational outcomes, December 2018

141 Tomasz **Makarewicz,** Traders, forecasters and financial instability: A model of individual learning of anchor-and-adjustment heuristics, January 2019

142 Carolin **Martin**, Noemi **Schmitt** and Frank **Westerhoff**, Housing markets, expectation formation and interest rates, January 2019

143 Marc P. **Saur**, Markus G. **Schlatterer** and Stefanie Y. **Schmitt**, Horizontal product differentiation with limited attentive consumers, January 2019

144 Tim **Hagenhoff** and Joep **Lustenhouwer**, The Rationality Bias, February 2019

145 Philipp **Mundt** and Ilfan **Oh**, Asymmetric competition, risk, and return distribution, February 2019

146 Ilfan **Oh**, Autonomy of Profit Rate Distribution and Its Dynamics from Firm Size Measures: A Statistical Equilibrium Approach, February 2019

147 Philipp **Mundt**, Simone **Alfarano** and Mishael **Milakovic**, Exploiting ergodicity in forecasts of corporate profitability, March 2019