Triebs, Thomas; Tumlinson, Justin

Conference Paper

Learning Capitalism the Hard Way—Evidence from Germany's Reunification

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Economic History, No. F13-V2

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Learning Capitalism the Hard Way—Evidence from Germany's Reunification

February 29, 2016

In a world where the future is uncertain and firms do not know the model, forecast ability matters. We ask whether, as predicted by rational expectations, forecast ability is uniform. And if not, whether firms learn. Firm level forecast ability is measured as forecast error based on the Ifo Institute's Business Climate Survey. We find that contrary to the prediction of rational expectations, forecast errors are persistent but we do not find any evidence that firms learn with age only. Then, we exploit German reunification, as a natural experiment where firms in the East are treated with ignorance about the state of the market. As predicted by our formal model of learning, firms in the East make larger forecast errors relative to the West. And over time this gap decreases. We argue that the initially higher forecast errors in the East are due to ignorance and not different market conditions. And, convergence is due to learning in the East, rather than convergence in market conditions.

PLEASE DO NOT CITE OR DISTRIBUTE

Keywords: Business forecasts; Firm behavior; Learning

1 Introduction

Forecasting matters in a world where the future is uncertain, just-in-time production (along the entire supply chain) is not feasible and firms need to plan and budget. Forecast errors have economic consequences, most obviously, in the form of input waste or foregone revenue. But economics generally assumes that all firms are equally capable forecasters and that forecast errors arise only from environmental surprises. The dominant theory of forecasting in the economic literature, *rational expectations*[1], leaves no room for variation in forecasting ability, much less learning. Firms always perfectly incorporate all available information to deliver minimum error forecasts; to the extent that some firms consistently forecast better than others, it is because they do so under more predictable conditions.

In this paper we ask two questions. (1) Are firms as uniform in their ability to forecast the future state of the market as rational expectations assumes? On the contrary, we find that the relative quality of firms' forecasts strongly predict the relative quality of their future forecasts, even among firms operating in the same industry, indicating differences in firms' ability to predict future market conditions. This observation opens a host of questions about the sources and consequences of such ability differences. We focus our investigation into this potentially broad and impactful arena by asking (2) can firms' ability to forecast change—in particular, can they learn?

Since the analog to human learning is natural, we first investigate firm age as a correlate of forecasting ability. Of course, as with individuals, disentangling the mechanisms of learning potentially associated with aging is non-trivial. Infant firms differ from mature ones in many ways besides forecasting ability: they are smaller, their employees tend to be younger, their markets tend to be newer and so on. Furthermore, the effects of these correlates of age may work in offsetting directions. Perhaps then, it is unsurprising that we find no general, systematic effect of age on the quality of firm forecasts of market conditions.

The ideal experiment to test the effects of experience on firm learning would be to exogenously place a cross-section of mature firms into a new market environment alongside experienced counterparts and compare their forecasts of their subsequently shared market conditions. The German Reunification of 1990 was just such an event. A homogeneous population was abruptly divided in 1949, and for four decades firms in East Germany operated under a master-planned, communist economy. For these firms of all sizes, maturities and across the spectrum

[1] In particular, Muth (1961) proposed the theory of rational expectations where the subjective probability distributions of outcomes matches the objective probability distribution of outcomes. With rational expectations, firms' forecast errors are independently and identically distributed and firms are never systematically wrong; they know the true model. Also, firms efficiently revise their forecasts using all available information.

of industries, market states were dictated, not predicted. Then suddenly, and quite unexpectedly, with the fall of the Berlin Wall on 9 November 1989, these firms were thrust into the free market economy of the West. Official institutions were unified rapidly and access to the global capitalist markets appeared suddenly. But did firms adapt as quickly as rational expectations predicted they would? Figure 1 plots forecast error magnitudes (ignoring whether firms are overoptimistic or overpessimistic) by region. Initially, Eastern firms made very large forecast errors compared to the West. Over time forecast errors in the East decrease and converge with forecast errors in the West. Our controlled regressions confirm the coarse implications of Figure 1: forecast errors across the two regions required two decades to converge (i.e. become empirically indistinguishable), despite the fact that actually market conditions converged very quickly.

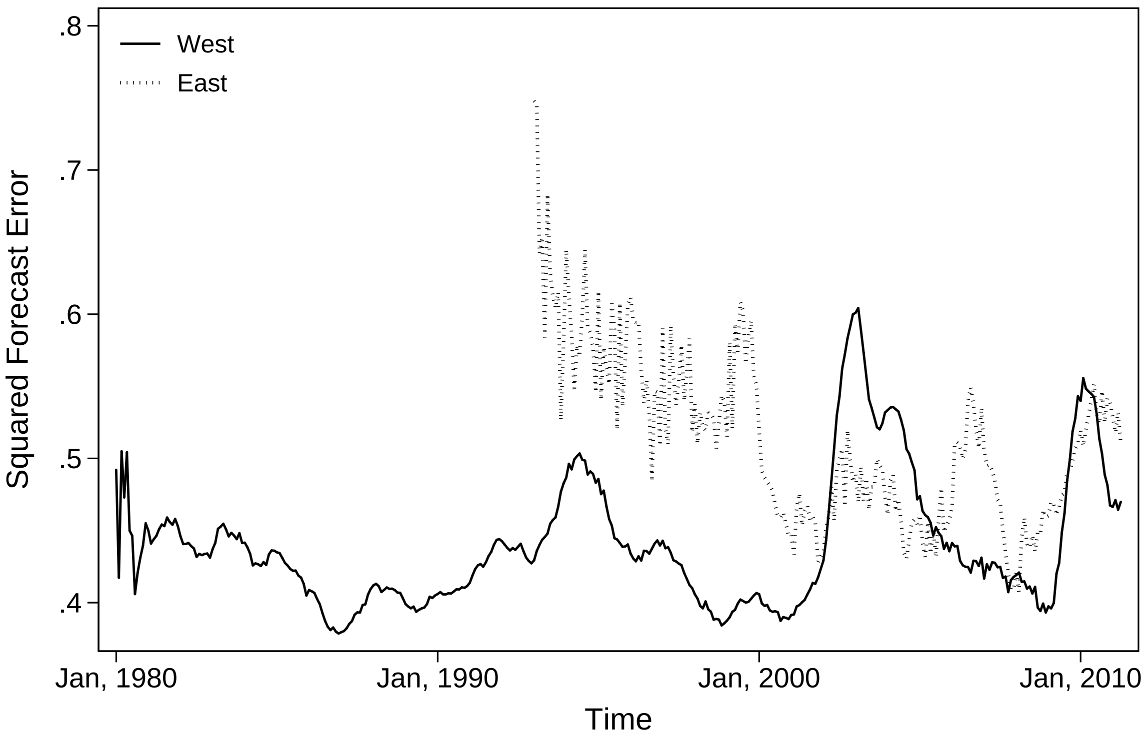


Figure 1: Forecast Error Magnitude

Notes: This graph plots the average forecast error magnitudes for East and West. The data is smoothed by averaging over 12 months. We only include firms that lived through Reunification in 1990.

To structure our analysis we introduce a formal model of (Bayesian) firm learning. In the model, the market state is drawn from a distribution whose parameters are initially unknown. Over time firms learn the model parameters. The model distinguishes between two types of market variance, distinctly observable to the econometrician: (1) volatility—the variance in realized market states—and, (2) uncertainty—disagreement among market players about the future state. We use the model to make empirical predictions about the error magnitudes and learning rates as a function of these market features.

The scarcity of data across a broad cross-section of firms has generally hampered analyses of forecasting at the firm level. However, we use a unique micro data set of the Ifo Institute to test our theoretical predictions. Every month the Ifo Institute's Business Climate Survey collects the near term expectations and assessment of the current market state from a large cross section of German manufacturing establishments.

Our study, although one of the first to examine market forecasting and learning at the firm level, is not without limitations. Since the primary purpose of the survey since 1949 has been to provide leading indicators of macro market health, large, established firms are oversampled. Thus, despite the fact that we do not broadly observe learning over age, we cannot rule out that startups very quickly reach a steady state. Furthermore, although we measure the learning of Eastern firms that live through Reunification, the mechanisms remain somewhat obscure. Given that our natural experiment shocked not just Eastern firms, but the individuals and non-firm institutions of East Germany quite radically, we cannot disentangle firm level learning from societal level learning.

Whether firms learn and if so at what rate is important for economic policy. During the recent Great Recession firms in several countries received subsidies to prevent their market exit. The rationale was to prevent the loss of knowledge and to bring production back swiftly after the recession. Forecast ability is a part of the firm knowledge that might be worth preserving if learning is slow or costly.

Furthermore, we live in period of geopolitical upheaval unseen since the collapse of the Iron Curtain and the Reunification of Germany. The lessons of this switch to democracy and capitalism may help set the correct expectations for the hypothetical reunifications of other countries like Korea or the political revolutions sweeping the Middle East—learning capitalism may take much longer than building its formal institutions.

The outline of this paper is as follows. Section 2 develops a formal model of production and learning. Section 3 introduces our empirical model and identification strategy. Section 4 introduces the data and measurement. Section 5 gives the results and section 6 concludes.

2 Theoretical Model

To structure our thinking we present stylized models of our data generating process under (1) pure rational expectations and (2) learning. The two models are identical, except that in the learning model, firms do not know one of the distributional parameters of the data generating process—they learn it over time. These simple models should not be taken as trying to capture the complexities of the true forecasting procedure that firms execute in predicting their future business conditions, but rather to parsimoniously focus on the intuitive difference between rational expectations and learning models.

At the beginning of period t, nature draws two hidden states relevant for firm i's change in business conditions S_{it} (literally *Geschäftslage*): X_{it} is an idiosyncratic firm state distributed $Normal\left(\mu_{X_i}, \sigma^2_{X_i}\right)$, and Y_t is a market state common to all firms distributed $Normal\left(\mu_Y, \sigma^2_Y\right)$, where X_{it} and Y_t are independent. For simplicity, assume $S_{it} = X_{it} + Y_t$. The information available at the beginning of of period t includes all *previous* state realizations $\Omega_{it} = \{X_{it-1}, \ldots, X_{i1}, Y_{t-1}, \ldots, Y_1\}$ and a firm specific signal of the current market state $\hat{Y}_{it} = Y_t + \varepsilon_{it}$, where white noise ε_{it} is, for all firms, distributed $Normal\left(0, \sigma^2_\varepsilon\right)$. Thus, $X_{it}|\,\hat{Y}_{it}, \Omega_{it} = X_{it}$ and $Y_t|\,\hat{Y}_{it}, \Omega_{it}$ are still independent. The key difference between the *rational expectations* model and the *learning model* is that in the latter, firms do not know the value of μ_Y; they must learn it. Under the learning model, firm i holds prior beliefs μ_{iY0} with prior variance σ^2_{iY0} about the mean of the market state variable, but these beliefs are updated over time.

Then, the firm makes a prediction about its state of business equal to the sum of conditional forecasts about its idiosyncratic and market states.

$$\bar{S}_{it} = \bar{X}_{it} + \bar{Y}_{it} = E\left[X_{it}\right] + E\left[Y_t|\,\hat{Y}_{it}, \Omega_{it}\right]$$

At the end of period t the realized state variables are revealed, and a forecast error $S_{it} - \bar{S}_{it}$ is computed. This error is directional. A positive value indicates that the firm was pessimistic—it predicted a worse change in business state than actually occurred. A negative value indicates that the firm was optimistic—it predicted a better change in state of business than actually occurred. We are interested in the average *magnitude* of this error or so-called *mean squared error*

$$\begin{aligned}
E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] &= E\left[\left((X_{it} + Y_t) - \left(E\left[X_{it}\right] + E\left[Y_t|\,\hat{Y}_{it}, \Omega_{it}\right]\right)\right)^2\right] \\
&= E\left[\left(X_{it} - E\left[X_{it}\right]\right)^2\right] + E\left[\left(Y_t - E\left[Y_t|\,\hat{Y}_{it}, \Omega_{it}\right]\right)^2\right] \\
&+ 2E\left[\left(X_{it} - E\left[X_{it}\right]\right)\left(Y_t - E\left[Y_t|\,\hat{Y}_{it}, \Omega_{it}\right]\right)\right] \\
&= Var\left[X_{it}\right] + Var\left[Y_t|\,\hat{Y}_{it}, \Omega_{it}\right]
\end{aligned}$$

where the independence of X_{it} and Y_t implies

$$2E\left[\left(X_{it}-E\left[X_{it}\right]\right)\left(Y_t-E\left[Y_t\right|\hat{Y}_{it},\Omega_{it}\right]\right)\right]=Cov\left(X_{it},Y_t\right|\hat{Y}_{it},\Omega_{it}\right)=0$$

The goal of the following analysis to compute comparative statics on the error magnitude with respect to the number of observations, the sample variance of the realizations over time and the sample variance of the predictions across firms.

2.1 Rational Expectations

When the distribution parameters of Y_t are known, history Ω_{it} is irrelevant to making predictions about the future, but current market signals $\hat{Y}_{it}$ are quite useful. It is well-known that the mean and variance of correct (normally distributed) posterior beliefs over Y_t are

$$\bar{Y}_{it}=E\left[Y_t\right|\hat{Y}_{it},\Omega_{it}\right]=\frac{\sigma_\varepsilon^2}{\sigma_Y^2+\sigma_\varepsilon^2}\mu_Y+\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\hat{Y}_{it} \tag{1}$$

$$\bar{\sigma}_{iy}^2=Var\left[Y_t\right|\hat{Y}_{it},\Omega_{it}\right]=\frac{\sigma_Y^2\sigma_\varepsilon^2}{\sigma_Y^2+\sigma_\varepsilon^2} \tag{2}$$

respectively. Thus,

$$E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=Var\left[X_{it}\right]+Var\left[Y_t\right|\hat{Y}_{it},\Omega_{it}\right]=\sigma_X^2+\frac{\sigma_Y^2\sigma_\varepsilon^2}{\sigma_Y^2+\sigma_\varepsilon^2} \tag{3}$$

In this case, we can generate the following comparative statics on MSE under rational expectations: (1) the expected magnitude of forecast errors does not change over time,

$$\frac{d}{dt}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=0 \tag{4}$$

(2) the expected magnitude of forecast errors increases with the volatility of the market,

$$\frac{d}{d\sigma_Y^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=\left(\frac{\sigma_\varepsilon^2}{\sigma_Y^2+\sigma_\varepsilon^2}\right)^2>0 \tag{5}$$

and (3) the expected magnitude of forecast errors increases with the uncertainty of the market (equivalently the noisiness of market signals),

$$\frac{d}{d\sigma_\varepsilon^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=\left(\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\right)^2>0 \tag{6}$$

2.2 Learning

It is well known that the posterior predictive distribution of Y_t, normally distributed with unknown mean and variance σ_Y^2, *unconditional on the signal* is $Normal\left(\mu_{Yt}, \sigma_{iYt}^2 + \sigma_Y^2\right)$ where

$$\mu_{iYt} = \frac{\sigma_Y^2}{t\sigma_{iY0}^2 + \sigma_Y^2}\mu_{i0} + \frac{t\sigma_{iY0}^2}{t\sigma_{iY0}^2 + \sigma^2}\bar{Y}$$
$$\sigma_{iYt}^2 = \frac{\sigma_{iY0}^2\sigma_Y^2}{t\sigma_{iY0}^2 + \sigma_Y^2}$$

$\bar{Y}$ is the sample mean of realized market states up to time t. Thus, substituting μ_{Yt} and $\sigma_{iYt}^2 + \sigma_Y^2$ for μ_Y and σ_Y^2 in equations (1) and (2) of the previous subsection, the mean and variance of correct (normally distributed) posterior predictives over Y_t are

$$\bar{Y}_{it} = E\left[Y_t|\,\hat{Y}_{it}\right] = \frac{\sigma_\varepsilon^2}{\left(\sigma_{iYt}^2 + \sigma_Y^2\right) + \sigma_\varepsilon^2}\mu_{Yt} + \frac{\left(\sigma_{iYt}^2 + \sigma_Y^2\right)}{\left(\sigma_{iYt}^2 + \sigma_Y^2\right) + \sigma_\varepsilon^2}\hat{Y}_{it}$$
$$\bar{\sigma}_{iy}^2 = Var\left[Y_t|\,\hat{Y}_{it}\right] = \frac{\left(\sigma_{iYt}^2 + \sigma_Y^2\right)\sigma_\varepsilon^2}{\left(\sigma_{iYt}^2 + \sigma_Y^2\right) + \sigma_\varepsilon^2}$$

respectively. Thus, we can generate the following comparative statics on the MSE under learning by substituting $\sigma_{iYt}^2 + \sigma_Y^2$ for σ_Y^2 into eqn. (3) and taking derivatives. (1) The expected magnitude of forecast errors decreases over time,

$$\frac{d}{dt}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \frac{d}{dt}\frac{\left(\sigma_{iYt}^2 + \sigma_Y^2\right)\sigma_\varepsilon^2}{\left(\sigma_{iYt}^2 + \sigma_Y^2\right) + \sigma_\varepsilon^2} = -\sigma_Y^2\left(\frac{\sigma_{iY0}^2\sigma_\varepsilon^2}{t\sigma_{iY0}^2\left(\sigma_Y^2 + \sigma_\varepsilon^2\right) + \sigma_Y^2\left(\sigma_{iY0}^2 + \sigma_Y^2 + \sigma_\varepsilon^2\right)}\right)^2 < 0 \tag{7}$$

(2) The expected magnitude of forecast errors increases with market volatility,

$$\frac{d}{d\sigma_Y^2}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \frac{d}{d\sigma_Y^2}\frac{\left(\sigma_{iYt}^2 + \sigma_Y^2\right)\sigma_\varepsilon^2}{\left(\sigma_{iYt}^2 + \sigma_Y^2\right) + \sigma_\varepsilon^2} = \frac{\left(\left(t\sigma_{iY0}^2 + \sigma_Y^2\right)\sigma_\varepsilon^2\right)^2 + t\left(\sigma_{iY0}^2\sigma_\varepsilon^2\right)^2}{\left(t\sigma_{iY0}^2\left(\sigma_Y^2 + \sigma_\varepsilon^2\right) + \sigma_Y^2\left(\sigma_{iY0}^2 + \sigma_Y^2 + \sigma_\varepsilon^2\right)\right)^2} > 0 \tag{8}$$

(3) The expected magnitude of forecast errors increases with market uncertainty,

$$\frac{d}{d\sigma_\varepsilon^2}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \frac{d}{d\sigma_\varepsilon^2}\frac{\left(\sigma_{iYt}^2 + \sigma_Y^2\right)\sigma_\varepsilon^2}{\left(\sigma_{iYt}^2 + \sigma_Y^2\right) + \sigma_\varepsilon^2} = \left(\frac{\left(t\sigma_{iY0}^2 + \sigma_{iY0}^2 + \sigma_Y^2\right)\sigma_Y^2}{t\sigma_{iY0}^2\left(\sigma_Y^2 + \sigma_\varepsilon^2\right) + \sigma_Y^2\left(\sigma_{iY0}^2 + \sigma_Y^2 + \sigma_\varepsilon^2\right)}\right)^2 > 0 \tag{9}$$

Observe that these comparative statics converge precisely to those under rational expectations as $t \to \infty$. Also note that, by taking the derivative of these expressions with respect to time we can compute how learning evolves. (1) Learning slows over time:

$$\frac{d^2}{dt^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=\frac{2\sigma_{iY0}^2\sigma_Y^2\left(\sigma_Y^2+\sigma_\varepsilon^2\right)\left(\sigma_{iY0}^2\sigma_\varepsilon^2\right)^2}{\left(t\sigma_{iY0}^2\left(\sigma_Y^2+\sigma_\varepsilon^2\right)+\sigma_Y^2\left(\sigma_{iY0}^2+\sigma_Y^2+\sigma_\varepsilon^2\right)\right)^3}>0$$

(2) Although readily computed, the impact of market volatility on learning rate is ambiguous:

$$\frac{d^2}{dtd\sigma_Y^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=\frac{\left(t\sigma_{iY0}^2\left(\sigma_Y^2-\sigma_\varepsilon^2\right)+\sigma_Y^2\left(\sigma_{iY0}^2+3\sigma_Y^2+\sigma_\varepsilon^2\right)\right)\left(\sigma_{iY0}^2\sigma_\varepsilon^2\right)^2}{\left(t\sigma_{iY0}^2\left(\sigma_Y^2+\sigma_\varepsilon^2\right)+\sigma_Y^2\left(\sigma_{iY0}^2+\sigma_Y^2+\sigma_\varepsilon^2\right)\right)^3}$$

(3) The magnitude of forecast errors diminish more quickly in markets with high uncertainty:

$$\frac{d^2}{dtd\sigma_\varepsilon^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=-\frac{2\left(t\sigma_{iY0}^2+\sigma_{iY0}^2+\sigma_Y^2\right)\sigma_\varepsilon^2\left(\sigma_{iY0}^2\sigma_Y^2\right)^2}{\left(t\sigma_{iY0}^2\left(\sigma_Y^2+\sigma_\varepsilon^2\right)+\sigma_Y^2\left(\sigma_{iY0}^2+\sigma_Y^2+\sigma_\varepsilon^2\right)\right)^3}<0$$

The intuition for $\frac{d^2}{dtd\sigma_\varepsilon^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]<0$ is somewhat subtle. Early on a firm must rely heavily on the signal of the current market state to forecast. Thus, early on, firms in industries with poor quality signals do very poorly relative to those in industries with better signals. However, as time progresses, firms learn the mean of the market state distribution, and thus, become less reliant on the signals alone. Hence this learning is more important to firms in industries with noisy signals.

3 Data

3.1 Forecast Errors

Our data is from the Ifo Institute's business climate survey, which, to our knowledge, is the oldest survey on firm expectations and realizations in existence. It started in November 1949 for the Federal Republic of Germany and included the former East Germany beginning in 1991. Our sample is for months from 1980 to 2010 for West Germany and 1992 to 2010 for East Germany; we drop 1991 observations for the East, because administrative difficulties render these earliest Eastern observations unreliable. The data is at the product level. Initially, in 1980 we have more than 4000 products in the cross-section. At the end, in 2010 about 2500 products remain in our sample. The data is for manufacturing firms only. Following (Nerlove, 1983, footnote 15) we treat product-level observations as independent; there are very few multi-product firms. Firms enter, exit and occasionally do not respond to the survey over our sample period. Thus, the panel is unbalanced. We drop all firms that report only once (forecast errors for such firms clearly cannot be computed). Like most surveys of this kind, the responses are granular, in our case trichotomous (*i.e.* '+', '=' or '-'). As the response categories are ordered we can compare forecasts and realizations to construct forecast errors. We calculate forecast errors for the market state (*Geschäftslage*). For the market state the forecast question reads:

> Expectations for the next six months: Our market state for product X will be rather better, about the same, rather worse.

And the question about the realization reads:

> Current situation: We assess our market state for product X to be good, satisfactory, bad.

Note that the question about the realization relates to levels and not changes. Also, a mismatch between the time horizons for the questions about forecasts and realizations complicates forecast error calculation. We adopt the following solution.

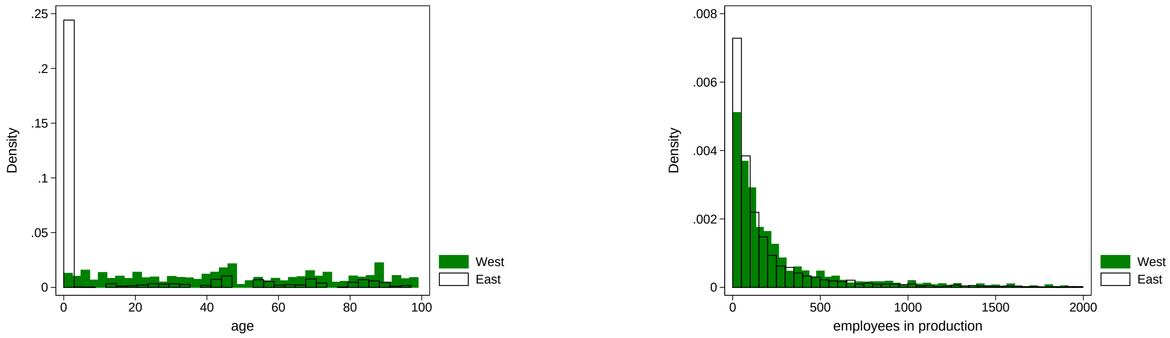
1. First, for every firm in every month we calculate a realized production measure equal to the unweighted six month average over realized firm level market states. We exclude months where there is no response.

2. The forecast error is calculated as that realization minus the forecast.

Last, we measure error sizes or magnitudes by taking the squares of the forecast errors. This inflates large directional errors but should not make any difference to our results. We observe a large number of neutral expectations. The error magnitude takes values between 1 and 4. On average firms give a neutral expectation just above 50 per cent of the time they respond. Given the near normal distribution of the firm level fractions of neutral responses we have no reason to believe that low forecasts errors are artifacts. However, the probability of giving a neutral expectation is not random. A linear probability model with firm fixed effects reveals that firms in the East are more likely to give a neutral expectations (higher average fixed effects), as are older firms, smaller firms, and first time respondents.

3.2 Firm Age

In our theory firms learn with experience. A natural proxy for experience is firm age. Besides the fact that age correlates with unobserved heterogeneity (e.g. management ability) our data on age has several shortcomings. In the West, the median firm age is 61 years and the median age of first time respondents is 51 years. We believe there are two reasons. First, because we collected the foundation date variable independently of the survey and we only collected it recently there is survivor bias. Second, it is possible that the survey only targets mature firms. In the East the problem is that a large proportion of firms is born at Reunification, which we believe is an artifact of privatization and does not reflect the actual operational experience of firms in the East (including the time before Reunification). Evidence for our hypothesis that for Eastern firms the recorded age is not their actual age is the fact that firm size variation right after Reunification in 1992 is not very different between East and West (Figure 2, right panel) but that the distributions for age are very different. In the East there is a large proportion of very young firm in 1992 (left panel).



(a) Age

(b) Size

Figure 2: Market Variability

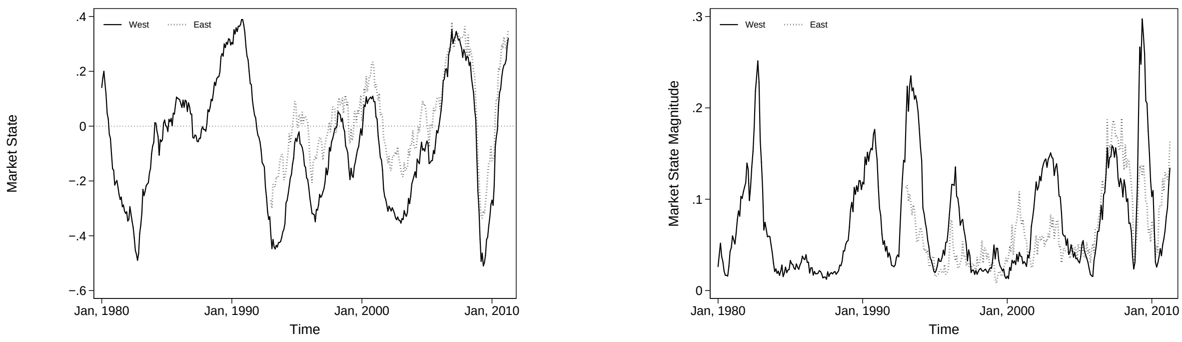
Notes: The left panel plots the densities of age for East and West in 1992. The right panel plots the densities of firm size (number of employees) for East and West in 1992.

3.3 Market Attributes

We characterize the market by the following attributes: its current state, its volatility, and its uncertainty. First, the state. In a "good" market firms do better than in a "poor" market. But generally, firms can only take advantage of a good market if they recognize and predict it as such. And they can only avoid the harm done by a poor market if they recognize (and forecast) it. The current state of a market is measured as the industry-region-time average response on the survey question about the market state, i.e. the market can be good, satisfactory, or bad. Markets are defined as industries at the two-digit level. The monthly average market states across industries for East and West are depicted in the left panel of Figure 3. The right panel also gives the square of the market state (a control variable in our regressions below), which indicates whether the market is neutral or extreme.

The variability of the market and the fact that the future is uncertain give rise to two additional attributes of the market: volatility and uncertainty. Volatility describes the range of possible states of the market. An analogy is the weather. Some regions have more volatile weather than others, e.g. the weather in Arizona is not very volatile. Ideally, we would measure current volatility as the variance over all possible states (real and imaginary) that the market can assume. But of course in a cross-section only a single (the realized) market state is observed and not its counterfactuals. Therefore, we measure volatility as the variance over first-differences of market states over the past two years. We take first-differences to control for trends. Figure 4 plots volatility in the left panel. We observe that generally volatility is higher in times of economic crisis, the range of potential market states expands.

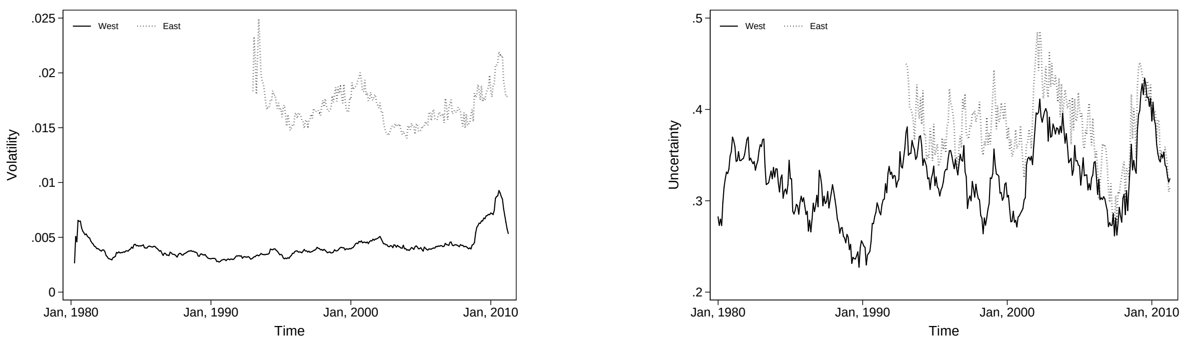
Besides the range of possible market states there is uncertainty about these states. Uncertainty relates to how easy it is to predict market states. Volatility and uncertainty are not independent. In a region where the weather is not volatile, it is easier to predict it. We measure uncertainty as the variance over firm-level forecasts for the market state in a given industry. This variance captures disagreement about the future market state. Uncertainty is depicted in the right panel of Figure 4.



(a) Market State **(b)** Market State Magnitude

Figure 3: Market State

Notes: These graphs plot the regional averages for the market state (left) panel and its square (right panel).



(a) Market Volatility **(b)** Market Uncertainty

Figure 4: Market Variability

Notes: The left panel plots average volatility across industries and regions. Volatility is measure as the variance over the first difference in market states for the past two years. The right panel plots average uncertainty by region. Uncertainty is the cross-sectional variance of the expectations within an industry.

Below we will argue that the convergence of forecast errors between East and West is due to learning in the East and not a convergence in market states. The aggregate descriptives here show that there is a level difference between the markets states across the regions but the medium-term cyclical patterns are similar. Also, in the East volatility is much higher; and there seems to be no convergence. Uncertainty is higher in the East initially but converges. In our regressions below we will control for the industry-level market characteristics. Table 1 gives summary statistics by region.

Table 1: Summary Statistics

	West				East			
	mean	*sd*	*max*	*min*	*mean*	*sd*	*max*	*min*
Squared Forecast Error	0.4743	0.704	4.00	0.00	0.6116	0.873	4.00	0.00
Market Volatility	0.0040	0.005	0.42	0.00	0.0188	0.042	1.64	0.00
Market Uncertainty	0.3195	0.061	1.07	0.00	0.3939	0.115	2.00	0.00
Squared Market State	0.0728	0.094	1.00	0.00	0.0778	0.119	1.00	0.00
Firm constrained (=1)	0.3484	0.476	1.00	0.00	0.5040	0.500	1.00	0.00
# Employees (thd.)	458.2627	2671.901	120000.00	0.00	182.3886	516.843	25000.00	0.00
Firm age	70.1689	54.682	404.00	0.00	34.1194	48.867	301.00	0.00
Exporter (=1)	0.7734	0.419	1.00	0.00	0.5684	0.495	1.00	0.00
Observations	1378840				243195			

4 Empirical model and Identification

4.1 Firm level ability and learning

First, we want to investigate whether firm level forecast ability actually varies and whether firms learn. We suggest that persistence in performance is evidence for different abilities. Our performance measure is a firm's forecast error percentile rank in a given industry and year. We average a firm's monthly forecast errors to obtain its annual error. We then estimate the following model

$$
\begin{aligned}
\text{RSqErr}_{ijt} \quad = \quad & \alpha_0 + \beta_1 \text{RSqErr}_{ijt-1} + \beta_2 \text{RSqErr}_{ijt-5} + \beta_3 \text{RSqErr}_{ijt-10} \\
& \gamma_1 \text{ SqMktState}_{jt} + \gamma_2 \text{ MktVol}_{jt} + \gamma_3 \text{ MktUncer}_{jt} + \gamma_4 \ln(\text{Empl}_{it}) + \gamma_5 \text{ Exp}_{it}
\end{aligned} \tag{10}
$$

where the percentile rank of the squared forecast error is a function of its lagged values for 1,5, and 10 years. Throughout, firms are indexed by i, industries by j, years by t, and month-years by m. Control variables are the squared market state[2], market volatility, market uncertainty, firm size as measured by the natural log of the number of employees in production and an indicator for whether the firm exports. We also effectively control for industry specific effects by taking the rank within an industry. Errors are clustered at the firm level. We estimate this model for Western firm only because we want to learn about persistence in general and we have longer series for Western firms.

If forecasting ability differs, ability might be given or firms can learn. One possibility is that firms learn over age. To test this we estimate the following model

$$
\begin{aligned}
\text{SqErr}_{ijm} \quad = \quad & \alpha_{\text{i}} + \delta_m + \beta_1 \text{Age}_{it} + \beta_2 \text{SqAge}_{it} \\
& \gamma_1 \text{ SqMktState}_{jm} + \gamma_2 \text{ MktVol}_{jm} + \gamma_3 \text{ MktUncer}_{jt} + \gamma_4 \ln(\text{Empl}_{it})
\end{aligned} \tag{11}
$$

where the squared forecast error is a function of the firm's age and age squared. The control variables are the same as above. Additionally we include firm and month-year fixed effects. Errors are clustered at the firm level. Again, we estimate the model for Western firms only. the age variable is subject to the limitations discussed above.

[2] When the dependent variable is directionless the explanatory variable should also be directionless. We capture the effects of extreme as opposed to neutral market states.

4.2 Reunification and learning

Firms might not only learn with age but also after abrupt changes to their market environment. Our next model takes advantage of the natural experiment of German reunification and asks whether firms that formerly operated in a planned economy learn the new market economy. The model is

$$
\begin{aligned}
\text{SqErr}_{ijm} \quad = \quad & \alpha_j + \delta_m + \beta_1 \text{East}_i + \beta_2 \text{t}_m + \beta_2 \text{t}_m^2 + \beta_3 \text{East}_i \times t_m + \beta_4 \text{East}_i \times \text{t}_m^2 \qquad (12) \\
& + \gamma_1 \ \text{SqMktState}_{jm} + \gamma_2 \ \text{MktVol}_{jm} + \gamma_3 \ \text{MktVol}_{jm} \times \text{t}_t + \gamma_4 \ \text{MktVol}_{jm} \times \text{t}_m \times \text{East}_i \\
& + \gamma_5 \ \text{MktUncer}_{jm} + \gamma_6 \ \text{MktVol}_{jm} \times \text{t}_t + \gamma_7 \ \text{MktVol}_{jm} \times \text{t}_m \times \text{East}_i \\
& + \gamma_8 \ \text{FrmState}_{im} + \gamma_9 \ \ln(\text{Empl}_{it}) + \gamma_{10} \text{Age}_{it} + \gamma_{11} \ \text{Exp}_{im}
\end{aligned}
$$

The dependent variable is the firm level squared forecast error. The first row in (12) contains the main variables of interest. Our theory above predicts that the average firm in the East makes larger forecast errors than in the West ($\beta_1 > 0$). It also predicts that East and West converge (at a diminishing rate). Given higher forecast errors in the East initially this implies $\beta_3 < 0$ and $\beta_4 > 0$. In the East the time trend t captures the common experience of firms after reunification. We make no predictions about the time trend in the West. Note that β_2 is dropped from the actual regression because it is collinear.

The second row has the current market state magnitude (SqMktState) as a control as well as market volatility (MktVol) and its interactions. Our theory includes γ_4 but makes no unambiguous prediction for its sign. The third row has market uncertainty (MktUncer) and its interactions. Our theory predicts that $\gamma_7 < 0$, learning is faster with higher uncertainty. The fourth row has firm level control variables: an indicator if the firm self-reports production constrains, the log number of employees, the firm's age and an indicator for whether the firm exports or not. We also control for industry and month-year fixed effects. Note that all trend variables at the month-year level are divided by 12 to facilitate the interpretation of the estimated coefficients below. We also estimate the model for exporters only. In a further specification we replace all the firm level control variables and the markets state magnitude control by a single proxy: the square of the firm level realisation, which we use to calculate the forecast error. This variable should capture everything that is of importance to the production of the firm.

To make sure that we only include firms that lived through reunification we restrict the sample as follows. For the West we only include firms whose first survey response was before 1987 and whose last response was after 1995. We do not have survey responses for Eastern firms before reunification so for them the requirement is to be born before 1987 and to have a response after 1995.

Identification of the coefficients is in the cross-section using an OLS estimator. The inclusion of an indicator variable for East precludes within identification. Errors are clustered at the firm level.

5 Results

First, we test whether forecast ability differs across firms. Table 2 presents our regression results for the model in (10). Column 1 omits controls. The controls added in the second column do not change the results or are individually significant statistically. The 1, 5, and 10 year lagged ranks of forecast quality are strong predictors for the current percentile rank. If a firm's percentile rank last year was 0.1 points higher it is still 0.047 points higher today. We suggest this is evidence that forecast ability differs across firms. Since ranks are computed within industry, the persistence of forecast quality cannot be explained by the relative ease of predicting one industry versus another; however, we acknowledge that even within industry, the individual businesses of some firms may be more predictable than others, but we cannot disentangle this from our data.

If ability differs, firms might learn to forecast well. Table 3 gives the results for model 11. We find no evidence that firms (in the West) learn with age. A caveat is that we do not have a representative sample of firm ages. In particular, our founding date reflects a sample bias: firm founding dates were collected for a subsample of firms recently answering the survey. Thus, failed firms, being more likely to have dropped out of the survey by the time recent waves occurred, will disproportionately not have associated founding dates in our sample. Of course, to the extent that one believes that forecasting ability is positively correlated to performance, this would lead to an overestimate of the effect of age on forecast quality. Even with this potential bias, we do not identify a significant effect of firm age on forecast quality.

The potential explanations for this non-result are manifold. Theoretically, age correlates with many things that may adversely influence forecast error, even if learning occurs: number of employees, general business complexity including market breadth and supply chain depth, and so on. Practically, our data suffers a relative shortage of very young and small firms, because the survey administrators wished to capture as much macro level information per respondent as possible. Hence, if learning occurs only over a very brief period beginning at inception, we may miss the statistical significance of this. So far we have found evidence that forecast ability differs across firms but not that firms learn with age.

Table 2: Forecast ability

	(1) No Controls	(2) Controls
L1.RSqError	0.4722***	0.4721***
	[0.000]	[0.000]
L5.RSqError	0.0881***	0.0881***
	[0.000]	[0.000]
L10.RSqError	0.0518***	0.0518***
	[0.000]	[0.000]
Market State^2		0.0187
		[0.157]
Uncertainty		0.0103
		[0.674]
Mkt. Volatility		0.3424
		[0.263]
log(Production Employees)		0.0003
		[0.754]
Exporter (=1)		0.0024
		[0.511]
Constant	0.1904***	0.1809***
	[0.000]	[0.000]
Observations	51714	51711

p-values in brackets. Errors are clustered at the firm level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Learning with age

	(1) All
Firm Age	0.0000
	[0.986]
Firm Age^2	-0.0000
	[0.933]
Mkt. Uncertainty	0.5730***
	[0.000]
Mkt. Uncertainty $\times$ Firm Age	0.0005
	[0.456]
Mkt. Volatility	-0.2248
	[0.873]
Mkt. Volatility $\times$ Firm Age	-0.0133
	[0.350]
Market State^2	0.2567***
	[0.000]
log(Production Employees)	-0.0008
	[0.900]
Constant	0.2417*
	[0.041]
Firm Fixed Effects	Yes
Month-Year Fixed Effects	Yes
Observations	582804

p-values in brackets. Errors are clustered at the firm level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In order to overcome the aforementioned limitations of age as driver of learning we exploit a natural shock affecting relevant firm level knowledge for a subsample of firms both old and young alike. When firms located in the East suddenly found themselves (and their local customers) operating in the global free market, the data generating process (DGP) which delivered market state—both supply and demand—suddenly changed, and of course, the distributional parameters governing this DGP were not announced. While firms and their management, no doubt, had prior beliefs about the DGP, it is not unreasonable to conjecture that these would be refined with experience.

Table 4 reports the results for the evolution of forecast errors after the East is treated with this market DGP reset, or "ignorance." The first column controls for market features, in particular the market state magnitude, volatility, and uncertainty. The second column additionally controls for firm level characteristics. Since, exporters in both East and West are more likely to be serving the same global demand than those manufacturing for the domestic German market, we restrict our sample to exporters in the third column. In the fourth column, to isolate the effects of forecasts from idiosyncratic firm shocks, we control for the actual *ex post* realisation of business state reported by the firm.

We find broad support for the predictions of the theory: (1) Eastern firms, shocked with market "ignorance" forecast worse. (2.a & b) Western forecast quality holds stable (except in column 4 where forecast errors in the West decrease over time), while it improves, at a diminishing rate, in the East. Convergence in error size occurs more gradually than one might expect, given how quickly the distribution of actual market states converged (more on this in the next section). Back of the envelope calculations with our reported coefficients for east and east-time interaction (ignoring all other coefficients) suggests that it took approximately two decades for Eastern firms to learn the market DGP as well as their Western counterparts. (3a & b) Our theory predicted that market volatility would lead to greater errors, but we identify no significant effect. The theory's prediction for market volatility's role in the rate of learning is ambiguous, and indeed we find no consistent or significant pattern in the volatility-time-east interaction in any of our specifications. (4a & b) We find strong support for the theory's prediction that uncertainty (or noise in firm level signals of the future market state) positively affects forecast errors. We generally also find support for its prediction that uncertainty increases the learning rate. Thus, the only prediction of the model not supported—but also not refuted—is that market volatility impedes forecast quality. In future work, we may be able to resolve this ambiguity by controlling for further correlates of market volatility.

Table 4: Reunification and learning

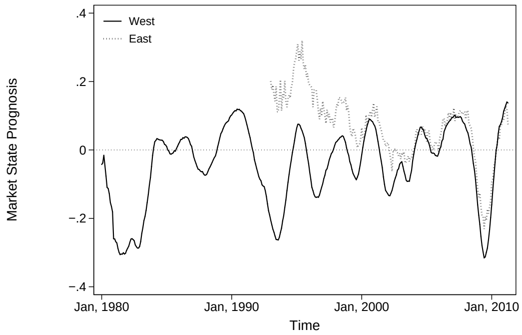
	(1) Market Controls	(2) & Firm Controls	(3) & Exporter only	(4) Realised Control
East (=1)	3.7739***	3.1720**	4.4547**	3.7863***
	[0.000]	[0.007]	[0.004]	[0.000]
Time	−0.0002	0.0009	0.0020	−0.0033*
	[0.909]	[0.685]	[0.439]	[0.028]
Time × East	−0.1681***	−0.1389*	−0.1945**	−0.1716***
	[0.001]	[0.013]	[0.008]	[0.000]
$\text{Time}^2 \times$ East	0.0019***	0.0016*	0.0022*	0.0020***
	[0.001]	[0.018]	[0.011]	[0.000]
Mkt. Volatility	−0.9344	−1.1558	−2.8863	−1.2971
	[0.629]	[0.584]	[0.310]	[0.444]
Mkt. Volatility × Time	−0.0267	0.0124	0.0562	−0.0216
	[0.693]	[0.869]	[0.546]	[0.711]
Mkt. Volatility × Time × East	0.0246	−0.0030	−0.0182	0.0267
	[0.473]	[0.947]	[0.754]	[0.370]
Mkt. Uncertainty	0.6143***	0.6175*	0.7205*	0.5251***
	[0.000]	[0.014]	[0.014]	[0.000]
Mkt. Uncertainty × Time	0.0028	0.0012	−0.0011	0.0042
	[0.523]	[0.861]	[0.884]	[0.292]
Mkt. Uncertainty × Time × East	−0.0059*	−0.0049	−0.0081*	−0.0028
	[0.019]	[0.101]	[0.032]	[0.232]
State Realisation^2				0.5708***
				[0.000]
Market State^2	0.3010***	0.2498***	0.2447***	
	[0.000]	[0.000]	[0.000]	
Firm State (Constrained =1)		0.1465***	0.1546***	
		[0.000]	[0.000]	
log(Production Employees)		0.0116**	0.0119**	
		[0.003]	[0.007]	
Exporter (=1)		−0.0004		
		[0.979]		
Firm Age		−0.0000	−0.0000	
		[0.738]	[0.840]	
Constant	0.2175***	0.0960	0.0202	0.1854***
	[0.000]	[0.324]	[0.855]	[0.001]
Industry Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Month-Year Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	845661	150030	118020	845661

p-values in brackets. Errors are clustered at the firm level.

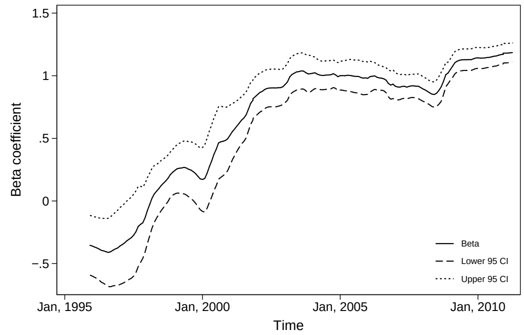
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.1 What drives the forecast error: expectation or realization?

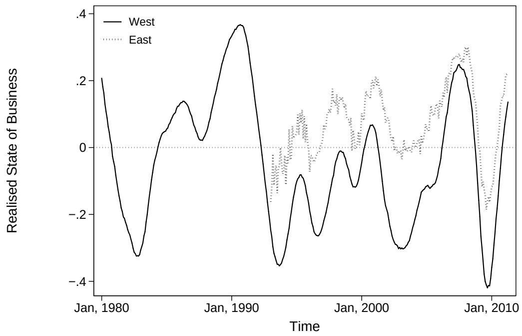
In the regressions in Table 4 we controlled for the different attributes of the market state because it is possible that the higher forecast errors in the East are due to different market characteristics rather than bad forecasting. An error always occurs as the result of a bad expectation for a given realisation, *i.e.* expectation is the decision variable. To provide more evidence that the higher forecast errors in the east are driven by bad expectations rather than different market circumstances we separate the forecast error into its components, expectation and realization, and correlate their aggregated time series. Figure 5 compares the time series expectations and realisation across the two regions. The left panels plot the raw series and the right panels give the beta coefficients (and the confidence intervals) from a regression of the time series for the West on the time series of the East. The estimated coefficient varies over time because we run the regressions for rolling 5 year windows. When we compare the two panels on the right we see that the correlation for the expectation gradually increases until about 2003. The correlation for the realisation increases much faster after reunification; then moves in cycles. We argue this is evidence that initially higher forecast errors in the East are due to ignorance, rather than different market states. And that the decline in Eastern forecast errors is due to learning, rather than changes in the market state.



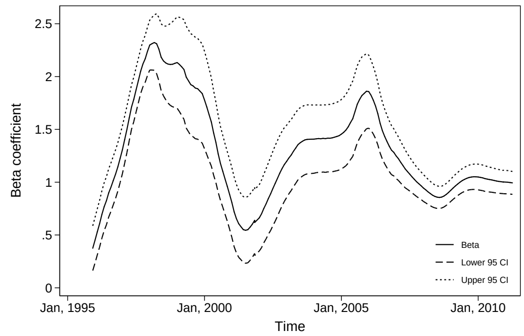
(a) Expectation



(b) Expectation



(c) Realisation



(d) Realisation

Figure 5: Components of the Market State Forecast Error

Notes: The left graphs plot the regional averages for the two components of the forecast error: prognosis and realisations. The data is smoothed by averaging over 12 months. We only include firms that lived through Reunification in 1990. The right graphs give the beta coefficients from the regression of the respective time series for the West on the series for the East after reunification using a rolling 5 year window.

6 Conclusion

A propos to our subject, in this initial investigation into firm level market forecasting, we confirmed several expectations and received a few surprises. We began by introducing a formal model of Bayesian learning, in which firms learn the distribution of market states, and thereby improve their forecasting ability. Although the dominant economic theory of forecasting, rational expectations, makes no explicit allowances for firm heterogeneity in forecasting ability or learning, our model of rational firm learning is a natural extension of this paradigm. This theory predicts that firm's forecasts improve with experience, but at an ever decreasing rate. The model predicts that both market volatility and market uncertainty make forecasting more difficult but that firms in uncertain markets reduce their forecast errors faster.

We find that forecast ability is persistent. However, we do not find that the theory's predictions over experience, when simple firm age is used as a proxy, are borne out. There are several potential explanations for this. Theoretically, many things besides experience vary with firm age, like workforce size and demographics, as well as product type and markets. The paucity of empirical research on firm level forecasts reflects the poor availability of data. Although the long-term, firm level survey of market conducted by the Ifo Institute is a rare exception, it also suffers some limitations. In particular, it does not contain a representative sample of the youngest firms. Thus, if forecasting ability is developed primarily in the infancy of firms, we may not be able to empirically observe it. Finally, it could be that firms, as organizations, separate from their employees, simply do not learn to forecast.

To overcome the theoretical and data limitations we face, we conduct an alternative evaluation of the theory. By comparing firms in former East and West Germany that survived their Reunification of Germany, we can test whether Eastern firms learn how to predict market states. They do. When time from Reunification is used as a proxy for experience, we find strong empirical support for all of the other predictions of the theory, except one: highly volatile markets do not significantly increase forecast errors—more work is needed to understand why. In particular, we find that forecast quality between Eastern and Western firms does eventually converge—after approximately two decades. Our evidence suggests that this is not due to convergence of the markets themselves, as these converge much quicker. Although the patterns we observe in the data are consistent with the theory, neither the theory nor our empirics can distinguish between institutional learning and the learning of eastern managers at the individual level.

Nevertheless our work, examining Germany, now one of the most potent economies in the world, tempers expectations for the political and economic upheavals occurring or contemplated around the world. From a potential unified Korea to the stabilization of the Arab Spring, it is reasonable to expect that learning capitalism will take much longer than building its formal institutions.

Of course, as is to be expected whenever fresh ground is broken, we have raised more questions about firm level forecasting than we have answered: What are the mechanisms by which firms learn? Besides learning, what else drives forecast heterogeneity between firms operating in similar environments? Does forecasting quality influence firm behavior or performance? Are there macro level consequences to firm heterogeneity in forecasting ability and responses to it? These are subjects for future investigation.

References

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, 29(3):315–335. ArticleType: research-article / Full publication date: Jul., 1961 / Copyright © 1961 The Econometric Society.

Nerlove, M. (1983). Expectations, plans, and realizations in theory and practice. *Econometrica*, 51(5):1251–1279. ArticleType: research-article / Full publication date: Sep., 1983 / Copyright © 1983 The Econometric Society.